

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Inc.

Gen protein - protein search using SW model

January 3, 2002, 10:00:20 : Search time 25.13 seconds
(without alignments) 1863.519 Million cell updates/sec

Title: US-09-722-602-10
Percent score: 34.25
Sequence: 1 AAFISALSRKRAVATDS.....LYNRSACVSVAFIEVTV 639

Scoring table:
BLOSUM62
Gap: 10.0, Export: 0.5

Searched: 522463 seqs, 7407290 residues
Total number of hits: 44,111,404
Maximum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match: 0%
Maximum Match: 100%
Listing first 45 summaries

Database: A.Geneseq_1101.*
1: /s1108/seqdata/genseq/genseq/AA1985.DAT.*
2: /s1108/seqdata/genseq/genseq/AA1986.DAT.*
3: /s1108/seqdata/genseq/genseq/AA1987.DAT.*
4: /s1108/seqdata/genseq/genseq/AA1988.DAT.*
5: /s1108/seqdata/genseq/genseq/AA1989.DAT.*
6: /s1108/seqdata/genseq/genseq/AA1990.DAT.*
7: /s1108/seqdata/genseq/genseq/AA1991.DAT.*
8: /s1108/seqdata/genseq/genseq/AA1992.DAT.*
9: /s1108/seqdata/genseq/genseq/AA1993.DAT.*
10: /s1108/seqdata/genseq/genseq/AA1994.DAT.*
11: /s1108/seqdata/genseq/genseq/AA1995.DAT.*
12: /s1108/seqdata/genseq/genseq/AA1996.DAT.*
13: /s1108/seqdata/genseq/genseq/AA1997.DAT.*
14: /s1108/seqdata/genseq/genseq/AA1998.DAT.*
15: /s1108/seqdata/genseq/genseq/AA1999.DAT.*
16: /s1108/seqdata/genseq/genseq/AA2000.DAT.*
17: /s1108/seqdata/genseq/genseq/AA2001.DAT.*
18: /s1108/seqdata/genseq/genseq/AA2002.DAT.*
19: /s1108/seqdata/genseq/genseq/AA2003.DAT.*
20: /s1108/seqdata/genseq/genseq/AA2004.DAT.*
21: /s1108/seqdata/genseq/genseq/AA2005.DAT.*
22: /s1108/seqdata/genseq/genseq/AA2006.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3420	99.9	680	22	AA063582	Amino acid sequence
2	3420	99.9	681	22	AA063583	Synthetic amino ac
3	2267	64.4	675	21	AA011372	Fusarium venenatum
4	2267	64.4	680	21	AA022795	Arabidopsis thalia
5	2267	64.4	681	21	AA022797	Arabidopsis thalia
6	2267	64.4	681	21	AA022799	Arabidopsis thalia
7	218.5	6.4	223	21	AA037599	Arabidopsis thalia
8	218.5	6.4	224	21	AA037599	Arabidopsis thalia
9	218.5	6.4	210	21	AA037597	Arabidopsis thalia
10	150	4.4	1146	20	AA035162	Chlamydia pneumoniae
11	150	4.4	1749	22	AA072314	Glutamic acid rich

12	142.5	4.2	521	22	AA075453	Human colon cancer
13	142.5	4.2	600	22	AA093443	Human protein sequ
14	142.5	4.2	642	22	AA067950	Human immune respo
15	141.5	4.1	699	20	AA008471	F. batistinum CP70
16	139.5	4.1	426	22	AA052793	Human protein sequ
17	138.5	4.0	748	22	AA040225	Human polypeptide
18	137.5	4.0	957	21	AA059288	Human MUC11 polype
19	137.5	4.0	957	22	AA024513	C9orf predicted am
20	137	4.0	336	22	AA063210	Human secreted pro
21	137	4.0	302	22	AA093372	Human protein sequ
22	136	4.0	424	22	AA094386	Human protein sequ
23	135.5	4.0	1638	20	AA001138	Enterococcus faeca
24	135.5	4.0	1638	20	AA001140	Enterococcus faeca
25	135.5	4.0	1638	20	AA001142	Enterococcus faeca
26	135	3.9	441	21	AA063208	Human secreted pro
27	134	3.9	441	21	AA063207	Gene 32 human secr
28	134	3.9	619	22	AA070388	Human host cell pr
29	133.5	3.9	631	21	AA059408	S. coelicolor Yaw
30	132.5	3.9	1186	22	AA050654	C. elegans APR-1 p
31	131.5	3.8	330	21	AA004587	Arabidopsis thalia
32	131.5	3.8	330	21	AA004586	Arabidopsis thalia
33	131.5	3.8	418	21	AA004585	Arabidopsis thalia
34	131.5	3.8	851	21	AA092718	V. furnissii extra
35	131	3.8	631	22	AA083274	Chlamydia trachoma
36	131	3.8	1358	17	AA096032	P. gingivalis bact
37	131	3.8	1358	19	AA069494	Haemagglutinin pro
38	130.5	3.8	1004	21	AA001841	Haemophilus influe
39	130.5	3.8	1010	21	AA001840	Haemophilus influe
40	130.5	3.8	1291	19	AA059912	Amino acid sequenc
41	130.5	3.8	2383	21	AA015945	E. coli proliferat
42	130	3.8	915	21	AA078514	Pullulanase amino
43	129	3.8	287	21	AA020134	Arabidopsis thalia
44	129	3.8	316	21	AA020133	Arabidopsis thalia
45	129	3.8	846	22	AA078417	A. taenialis-E. coli

ALIGNMENTS

RESULT 1
AA063582 standard; Protein: 680 AA.
ID AA063582:
AC AAC63582:
XX 15-OCT-2001 (first entry)
XX DT
XX CE Amino acid sequence of a galactose oxidase
XX CE
XX ME Methionine gamma-lyase; mda gene; free folding energy; gene shuffling,
KW directed evolution; molecular breeding; galactose oxidase.
XX OS Hypomyces rosellus.
XX OS
XX FN WC20015342-A2.
XX PD 02-AUG-2001.
XX PE 31-JAN-2001; 2001WO-US03186.
XX PP 31-JAN-2000; 2000US-0494921
XX PR 08-DEC-2000; 2000US-0734237.
XX (BIOC-) BIOFATATYTICS INC.
XX PA Kozzell DJ, Bui P, Hua L;
XX PT WPI: 2001-483235/52.
XX DR N-PSDB; AA074601.
XX PT Designing synthetic nucleic acid sequences for improved amplification,
XX PT expression in host cell, by comparing free energy of folding of a
XX PT starting polynucleotide and a modified polynucleotide having a codon

replacement -
claim 6: Page 111-117; 117pp; English.

The present sequence represents a galactose oxidase enzyme. The polynucleotide was modified using the method of the invention. The specification describes a method for designing a synthetic polynucleotide. The method comprises providing a starting polynucleotide, determining the predicted free energy of folding per base of the polynucleotide, modifying the polynucleotide by replacing a codon with a different codon to provide a modified polynucleotide, determining free energy of folding per base of the modified polynucleotide, and comparing this with that of the original polynucleotide. The method is useful for developing nucleic acid sequences that enhance expression of the encoded protein in a heterologous host. The design and preparation of the synthetic genes are used in application of gene shuffling, directed evolution and recombinant breeding methods. The method allows expression of genes from various organisms such as mammals, plants, yeast, fungi and bacteria in prokaryotic hosts, such as *Escherichia coli* and eukaryotic hosts at commercially viable levels, in particular proteins with low yield such as methionine gamma-lyase from *P. putida*.

Sequence: 680 AA.

Query: Rat L. 99.98; Score 3420; DB 22; Length 680;
Best Local Similarity 99.88; Pred. No. 1,3e-247;
Matches: 648; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 ASAGTSAISPNNAWVCTASAGSGNECNRAIKENKPTFWTFFVGNATGPPPHITITIMK 60
|||||
42 ASAGTSAISPNNAWVCTASAGSGNECNRAIKENKPTFWTFFVGNATGPPPHITITIMK 101
|||||
61 TTAGNGTSMPPGPNNQWGTGPPPVITSDGTNNGSPVYASGWSADSTKXSNETRP 120
|||||
102 TTAGNGTSMPPGPNNQWGTGPPPVITSDGTNNGSPVYASGWSADSTKXSNETRP 161
|||||
121 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 180
|||||
162 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 241
|||||
181 GAVLMSNNAWVCTASAGSGNECNRAIKENKPTFWTFFVGNATGPPPHITITIMK 240
|||||
222 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 281
|||||
241 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 300
|||||
282 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 341
|||||
341 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 360
|||||
342 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 401
|||||
361 GAVLMSNNAWVCTASAGSGNECNRAIKENKPTFWTFFVGNATGPPPHITITIMK 420
|||||
402 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 461
|||||
421 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 480
|||||
462 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 521
|||||
481 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 540
|||||
522 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 581
|||||
541 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 600
|||||
582 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 641
|||||
601 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 639
|||||
642 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 680

Result 2
AAG63583
ID AAG63583 standard; Protein; 681 AA.
XX
AC AAG63583;
XX
XX 15-OCT-2001 (first entry)
DT
XX
XX Synthetic amino acid sequence of a galactose oxidase.
XX
XX Methionine gamma-lyase; mda gene; free folding energy; gene shuffling;
XX directed evolution; molecular breeding; galactose oxidase.
XX
XX Hypomyces rosellus.
XX
XX W020015342-A2.
XX
XX 02 AUG 2001.
XX
XX 31-JAN-2001; 2001WO-0503186
XX
XX 31-JAN-2003; 2003US 0494921.
XX
XX 08-DEC-2003; 2003US 0734237.
XX
XX (BIOC) BIOCATALYTICS INC.
XX
XX Rozzell PJ, Hul P, Hua L;
XX
XX WPI, 2001 48335/52.
XX
XX N-PSDB; AAH74602.

Designing synthetic nucleic acid sequences for improved amplification, expression in host cells, by comparing free energy of folding of a starting polynucleotide and a modified polynucleotide having a codon replacement -

Claim 6: Page 115-117; 117pp; English

The present sequence represents a modified galactose oxidase enzyme. The polynucleotide was modified using the method of the invention. The specification describes a method for designing a synthetic polynucleotide. The method comprises providing a starting polynucleotide, determining the predicted free energy of folding per base of the polynucleotide, modifying the polynucleotide by replacing a codon with a different codon to provide a modified polynucleotide, determining free energy of folding per base of the modified polynucleotide, and comparing this with that of the original polynucleotide. The method is useful for developing nucleic acid sequences that enhance expression of the encoded protein in a heterologous host. The design and preparation of the synthetic genes are used in application of gene shuffling, directed evolution and molecular breeding methods. The method allows expression of genes from various organisms such as mammals, plants, yeast, fungi and bacteria in prokaryotic hosts, such as *Escherichia coli* and eukaryotic hosts at commercially viable levels, in particular proteins with low yield such as methionine gamma-lyase from *P. putida*.

Sequence: 681 AA;

Query: Match 99.98; Score 3420; DB 22; Length 681;
Best Local Similarity 99.88; Pred. No. 1,3e-247;
Matches: 648; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 ASAGTSAISPNNAWVCTASAGSGNECNRAIKENKPTFWTFFVGNATGPPPHITITIMK 60
|||||
43 ASAGTSAISPNNAWVCTASAGSGNECNRAIKENKPTFWTFFVGNATGPPPHITITIMK 102
|||||
61 TTAGNGTSMPPGPNNQWGTGPPPVITSDGTNNGSPVYASGWSADSTKXSNETRP 120
|||||
103 TTAGNGTSMPPGPNNQWGTGPPPVITSDGTNNGSPVYASGWSADSTKXSNETRP 162
|||||
121 AAVVPLVATTEANGCWTSIAETINVFQASSVTAPDGTGHWGPIIDPVPAAAIPTTS 180

XX 17-CT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO. 25862.
XX Protein identification: signal transduction pathway; metabolic pathway;
XX cDNA library: genetic mapping, gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX E1103345 Az.
XX 06 SEP-2000.
XX 25-FEB-2000; 2000EP 030139.
XX 25-FEB-1999; 9908-0121825.
XX 05-MAR-1999; 9908-0123180.
XX 09-MAR-1999; 9908-0123548.
XX 24-MAR-1999; 9908-0125788.
XX 25-MAR-1999; 9908-0126264.
XX 29-MAR-1999; 9908-0126785.
XX 01-APR-1999; 9908-0127462.
XX 06-APR-1999; 9908-0128234.
XX 08-APR-1999; 9908-0128714.
XX 16-APR-1999; 9908-0129845.
XX 19-APR-1999; 9908-0130077.
XX 21-APR-1999; 9908-0130449.
XX 24-APR-1999; 9908-0130510.
XX 28-APR-1999; 9908-0130641.
XX 30-APR-1999; 9908-0131449.
XX 04-MAY-1999; 9908-0132407.
XX 06-MAY-1999; 9908-0132485.
XX 06-MAY-1999; 9908-0132486.
XX 07-MAY-1999; 9908-0132487.
XX 11-MAY-1999; 9908-0132863.
XX 14-MAY-1999; 9908-0134256.
XX 14-MAY-1999; 9908-0134218.
XX 14-MAY-1999; 9908-0134219.
XX 14-MAY-1999; 9908-0134221.
XX 14-MAY-1999; 9908-0134370.
XX 18-MAY-1999; 9908-0134768.
XX 19-MAY-1999; 9908-0134941.
XX 20-MAY-1999; 9908-0135124.
XX 21-MAY-1999; 9908-0135353.
XX 24-MAY-1999; 9908-0135623.
XX 25-MAY-1999; 9908-0136021.
XX 27-MAY-1999; 9908-0136392.
XX 28-MAY-1999; 9908-0136782.
XX 01-JUN-1999; 9908-0137222.
XX 04-JUN-1999; 9908-0137528.
XX 07-JUN-1999; 9908-0137502.
XX 08-JUN-1999; 9908-0137722.
XX 10-JUN-1999; 9908-0138094.
XX 10-JUN-1999; 9908-0138540.
XX 10-JUN-1999; 9908-0138847.
XX 14-JUN-1999; 9908-0139119.
XX 16-JUN-1999; 9908-0139452.
XX 17-JUN-1999; 9908-0139463.
XX 17-JUN-1999; 9908-0139492.
XX 18-JUN-1999; 9908-0139454.
XX 18-JUN-1999; 9908-0139455.
XX 18-JUN-1999; 9908-0139456.
XX 18-JUN-1999; 9908-0139457.
XX 18-JUN-1999; 9908-0139458.
XX 18-JUN-1999; 9908-0139459.
XX 18-JUN-1999; 9908-0139460.
XX 18-JUN-1999; 9908-0139461.
XX 18-JUN-1999; 9908-0139462.
XX 18-JUN-1999; 9908-0139463.

PR 18-JUN-1999; 9908-0139750.
PR 18-JUN-1999; 9908-0139763.
PR 21-JUN-1999; 9908-0139817.
PR 22-JUN-1999; 9908-0139899.
PR 23-JUN-1999; 9908-0140353.
PR 23-JUN-1999; 9908-0140354.
PR 24-JUN-1999; 9908-0140595.
PR 28-JUN-1999; 9908-0140823.
PR 29-JUN-1999; 9908-0140991.
PR 30-JUN-1999; 9908-0141287.
PR 01-JUL-1999; 9908-0141842.
PR 01-JUL-1999; 9908-0142154.
PR 02-JUL-1999; 9908-0142055.
PR 06-JUL-1999; 9908-0142390.
PR 08-JUL-1999; 9908-0142803.
PR 09-JUL-1999; 9908-0142920.
PR 12-JUL-1999; 9908-0142927.
PR 13-JUL-1999; 9908-0143542.
PR 14-JUL-1999; 9908-0143624.
PR 15-JUL-1999; 9908-0144005.
PR 16-JUL-1999; 9908-0144085.
PR 16-JUL-1999; 9908-0144086.
PR 19-JUL-1999; 9908-0144325.
PR 19-JUL-1999; 9908-0144331.
PR 19-JUL-1999; 9908-0144332.
PR 19-JUL-1999; 9908-0144333.
PR 19-JUL-1999; 9908-0144334.
PR 19-JUL-1999; 9908-0144335.
PR 20-JUL-1999; 9908-0144352.
PR 20-JUL-1999; 9908-0144362.
PR 20-JUL-1999; 9908-0144384.
PR 21-JUL-1999; 9908-0144814.
PR 21-JUL-1999; 9908-0145086.
PR 21-JUL-1999; 9908-0145088.
PR 22-JUL-1999; 9908-0145085.
PR 22-JUL-1999; 9908-0145087.
PR 22-JUL-1999; 9908-0145089.
PR 22-JUL-1999; 9908-0145192.
PR 23-JUL-1999; 9908-0145145.
PR 23-JUL-1999; 9908-0145218.
PR 23-JUL-1999; 9908-0145224.
PR 24-JUL-1999; 9908-0145276.
PR 24-JUL-1999; 9908-0145413.
PR 27-JUL-1999; 9908-0145918.
PR 27-JUL-1999; 9908-0145919.
PR 28-JUL-1999; 9908-0145941.
PR 02-AUG-1999; 9908-0146386.
PR 02-AUG-1999; 9908-0146388.
PR 02-AUG-1999; 9908-0146389.
PR 03-AUG-1999; 9908-0147038.
PR 04-AUG-1999; 9908-0147204.
PR 04-AUG-1999; 9908-0147302.
PR 05-AUG-1999; 9908-0147302.
PR 05-AUG-1999; 9908-0147260.
PR 06-AUG-1999; 9908-0147303.
PR 06-AUG-1999; 9908-0147493.
PR 06-AUG-1999; 9908-0147493.
PR 09-AUG-1999; 9908-0147935.
PR 10-AUG-1999; 9908-0148171.
PR 11-AUG-1999; 9908-0148319.
PR 12-AUG-1999; 9908-0148341.
PR 13-AUG-1999; 9908-0148565.
PR 13-AUG-1999; 9908-0148684.
PR 13-AUG-1999; 9908-0149368.
PR 17-AUG-1999; 9908-0149175.
PR 18-AUG-1999; 9908-0149426.
PR 20-AUG-1999; 9908-0149422.
PR 20-AUG-1999; 9908-0149723.
PR 20-AUG-1999; 9908-0149929.
PR 23-AUG-1999; 9908-0149902.
PR 23-AUG-1999; 9908-0149930.
PR 25-AUG-1999; 9908-0150566.
PR 26-AUG-1999; 9908-0150884.

13 MAY-1999; 9905-0144218.
PR 13 MAY-1999; 9905-0134219.
PR 13 MAY-1999; 9905-0144221.
PR 13 MAY-1999; 9905-0134270.
PR 13 MAY-1999; 9905-0134478.
PR 13 MAY-1999; 9905-0134941.
PR 20 MAY-1999; 9905-0135124.
PR 21 MAY-1999; 9905-0135353.
PR 21 MAY-1999; 9905-0135629.
PR 25 MAY-1999; 9905-0136621.
PR 27 MAY-1999; 9905-0136392.
PR 28 MAY-1999; 9905-0136782.
PR 01 JUN-1999; 9905-0137222.
PR 04 JUN-1999; 9905-0137528.
PR 04 JUN-1999; 9905-0137502.
PR 07 JUN-1999; 9905-0137724.
PR 08 JUN-1999; 9905-0138094.
PR 10 JUN-1999; 9905-0138540.
PR 10 JUN-1999; 9905-0138847.
PR 14 JUN-1999; 9905-0139119.
PR 14 JUN-1999; 9905-0139452.
PR 16 JUN-1999; 9905-0139453.
PR 17 JUN-1999; 9905-0139492.
PR 18 JUN-1999; 9905-0139454.
PR 18 JUN-1999; 9905-0139455.
PR 18 JUN-1999; 9905-0139456.
PR 18 JUN-1999; 9905-0139457.
PR 18 JUN-1999; 9905-0139458.
PR 18 JUN-1999; 9905-0139459.
PR 18 JUN-1999; 9905-0139460.
PR 18 JUN-1999; 9905-0139461.
PR 18 JUN-1999; 9905-0139462.
PR 18 JUN-1999; 9905-0139463.
PR 18 JUN-1999; 9905-0139750.
PR 18 JUN-1999; 9905-0139763.
PR 21 JUN-1999; 9905-0139817.
PR 22 JUN-1999; 9905-0139889.
PR 24 JUN-1999; 9905-0140353.
PR 24 JUN-1999; 9905-0140354.
PR 24 JUN-1999; 9905-0140695.
PR 24 JUN-1999; 9905-0140823.
PR 24 JUN-1999; 9905-0140991.
PR 30 JUN-1999; 9905-0141287.
PR 01 JUL-1999; 9905-0141842.
PR 01 JUL-1999; 9905-0142154.
PR 02 JUL-1999; 9905-0142055.
PR 06 JUL-1999; 9905-0142390.
PR 08 JUL-1999; 9905-0142803.
PR 09 JUL-1999; 9905-0142920.
PR 12 JUL-1999; 9905-0142977.
PR 13 JUL-1999; 9905-0143542.
PR 14 JUL-1999; 9905-0143624.
PR 15 JUL-1999; 9905-0144605.
PR 15 JUL-1999; 9905-0144085.
PR 15 JUL-1999; 9905-0144325.
PR 15 JUL-1999; 9905-0144326.
PR 15 JUL-1999; 9905-0144331.
PR 15 JUL-1999; 9905-0144332.
PR 15 JUL-1999; 9905-0144333.
PR 15 JUL-1999; 9905-0144334.
PR 15 JUL-1999; 9905-0144335.
PR 15 JUL-1999; 9905-0144352.
PR 20 JUL-1999; 9905-0144632.
PR 20 JUL-1999; 9905-0144884.
PR 21 JUL-1999; 9905-0144814.
PR 21 JUL-1999; 9905-0145086.
PR 21 JUL-1999; 9905-0145087.
PR 22 JUL-1999; 9905-0145089.
PR 22 JUL-1999; 9905-0145192.
PR 24 JUL-1999; 9905-0145218.
PR 24 JUL-1999; 9905-0145218.

23 JUL-1999; 9905-0145224.
PR 26 JUL-1999; 9905-0145276.
PR 27 JUL-1999; 9905-0145913.
PR 27 JUL-1999; 9905-0145918.
PR 27 JUL-1999; 9905-0145919.
PR 28 JUL-1999; 9905-0145951.
PR 02 AUG-1999; 9905-0146386.
PR 02 AUG-1999; 9905-0146389.
PR 03 AUG-1999; 9905-0147038.
PR 04 AUG-1999; 9905-0147204.
PR 04 AUG-1999; 9905-0147305.
PR 05 AUG-1999; 9905-0147302.
PR 05 AUG-1999; 9905-0147303.
PR 05 AUG-1999; 9905-0147304.
PR 06 AUG-1999; 9905-0147416.
PR 06 AUG-1999; 9905-0147493.
PR 09 AUG-1999; 9905-0147935.
PR 10 AUG-1999; 9905-0148171.
PR 11 AUG-1999; 9905-0148319.
PR 12 AUG-1999; 9905-0148341.
PR 13 AUG-1999; 9905-0148565.
PR 13 AUG-1999; 9905-0148684.
PR 16 AUG-1999; 9905-0149368.
PR 17 AUG-1999; 9905-0149175.
PR 18 AUG-1999; 9905-0149426.
PR 20 AUG-1999; 9905-0149723.
PR 20 AUG-1999; 9905-0149929.
PR 23 AUG-1999; 9905-0149902.
PR 23 AUG-1999; 9905-0149930.
PR 25 AUG-1999; 9905-0150566.
PR 26 AUG-1999; 9905-0150884.
PR 27 AUG-1999; 9905-0151065.
PR 27 AUG-1999; 9905-0151066.
PR 27 AUG-1999; 9905-0151080.
PR 30 AUG-1999; 9905-0151303.
PR 31 AUG-1999; 9905-0151348.
PR 01 SEP-1999; 9905-0151930.
PR 07 SEP-1999; 9905-0152363.
PR 10 SEP-1999; 9905-0153070.
PR 13 SEP-1999; 9905-0153758.
PR 15 SEP-1999; 9905-0154018.
PR 16 SEP-1999; 9905-0154039.
PR 20 SEP-1999; 9905-0154779.
PR 22 SEP-1999; 9905-0155139.
PR 24 SEP-1999; 9905-0155486.
PR 24 SEP-1999; 9905-0155659.
PR 28 SEP-1999; 9905-0156458.
PR 29 SEP-1999; 9905-0156586.
PR 04 OCT-1999; 9905-0157117.
PR 05 OCT-1999; 9905-0157753.
PR 06 OCT-1999; 9905-0157865.
PR 07 OCT-1999; 9905-0158029.
PR 08 OCT-1999; 9905-0158232.
PR 12 OCT-1999; 9905-0158369.
PR 13 OCT-1999; 9905-0159293.
PR 13 OCT-1999; 9905-0159294.
PR 13 OCT-1999; 9905-0159295.
PR 14 OCT-1999; 9905-0159329.
PR 14 OCT-1999; 9905-0159330.
PR 14 OCT-1999; 9905-0159331.
PR 14 OCT-1999; 9905-0159637.
PR 14 OCT-1999; 9905-0159638.
PR 18 OCT-1999; 9905-0159584.
PR 18 OCT-1999; 9905-0160741.
PR 21 OCT-1999; 9905-0160757.
PR 21 OCT-1999; 9905-0160768.
PR 21 OCT-1999; 9905-0160770.
PR 21 OCT-1999; 9905-0160814.
PR 21 OCT-1999; 9905-0160815.
PR 22 OCT-1999; 9905-0160980.
PR 22 OCT-1999; 9905-0160981.

28-JUN-1999 9908-0140823.
PR 29-JUN-1999 9908-0140991.
PR 30-JUN-1999 9908-0141287.
PR 01-JUL-1999 9908-0141642.
PR 01-JUL-1999 9908-0142154.
PR 02-JUL-1999 9908-0142055.
PR 03-JUL-1999 9908-0142390.
PR 04-JUL-1999 9908-0142403.
PR 05-JUL-1999 9908-0142920.
PR 12-JUL-1999 9908-0142977.
PR 13-JUL-1999 9908-0143542.
PR 14-JUL-1999 9908-0143624.
PR 15-JUL-1999 9908-0144005.
PR 16-JUL-1999 9908-0144085.
PR 16-JUL-1999 9908-0144086.
PR 17-JUL-1999 9908-0144125.
PR 17-JUL-1999 9908-0144341.
PR 19-JUL-1999 9908-0144332.
PR 19-JUL-1999 9908-0144333.
PR 19-JUL-1999 9908-0144334.
PR 19-JUL-1999 9908-0144335.
PR 20-JUL-1999 9908-0144352.
PR 20-JUL-1999 9908-0144632.
PR 20-JUL-1999 9908-0144884.
PR 21-JUL-1999 9908-0144814.
PR 21-JUL-1999 9908-0145086.
PR 21-JUL-1999 9908-0145088.
PR 22-JUL-1999 9908-0145085.
PR 22-JUL-1999 9908-0145087.
PR 22-JUL-1999 9908-0145089.
PR 22-JUL-1999 9908-0145192.
PR 24-JUL-1999 9908-0145145.
PR 24-JUL-1999 9908-0145218.
PR 24-JUL-1999 9908-0145224.
PR 25-JUL-1999 9908-0145276.
PR 25-JUL-1999 9908-0145913.
PR 25-JUL-1999 9908-0145918.
PR 26-JUL-1999 9908-0145919.
PR 26-JUL-1999 9908-0145951.
PR 26-JUL-1999 9908-0146386.
PR 02-AUG-1999 9908-0146388.
PR 02-AUG-1999 9908-0146389.
PR 03-AUG-1999 9908-0147038.
PR 04-AUG-1999 9908-0147204.
PR 04-AUG-1999 9908-0147302.
PR 05-AUG-1999 9908-0147192.
PR 05-AUG-1999 9908-0147260.
PR 06-AUG-1999 9908-0147303.
PR 06-AUG-1999 9908-0147416.
PR 09-AUG-1999 9908-0147493.
PR 09-AUG-1999 9908-0147935.
PR 11-AUG-1999 9908-0148171.
PR 11-AUG-1999 9908-0148311.
PR 13-AUG-1999 9908-0148341.
PR 13-AUG-1999 9908-0148565.
PR 13-AUG-1999 9908-0148664.
PR 14-AUG-1999 9908-0149368.
PR 17-AUG-1999 9908-0149175.
PR 18-AUG-1999 9908-0149426.
PR 20-AUG-1999 9908-0149722.
PR 20-AUG-1999 9908-0149723.
PR 20-AUG-1999 9908-0149922.
PR 23-AUG-1999 9908-0149902.
PR 23-AUG-1999 9908-0149930.
PR 24-AUG-1999 9908-0150566.
PR 26-AUG-1999 9908-0150884.
PR 27-AUG-1999 9908-0151065.
PR 27-AUG-1999 9908-0151066.
PR 27-AUG-1999 9908-0151080.
PR 27-AUG-1999 9908-0151303.
PR 31-AUG-1999 9908-0151438.
PR 01-SEP-1999 9908-0151930.
PR 07-SEP-1999 9908-0152363.

PR 10-SEP-1999 9908-0153070.
PR 13-SEP-1999 9908-0153758.
PR 15-SEP-1999 9908-0154018.
PR 16-SEP-1999 9908-0154039.
PR 20-SEP-1999 9908-0154779.
PR 22-SEP-1999 9908-0155139.
PR 23-SEP-1999 9908-0155486.
PR 24-SEP-1999 9908-0155659.
PR 28-SEP-1999 9908-0156458.
PR 29-SEP-1999 9908-0156596.
PR 04-OCT-1999 9908-0157117.
PR 05-OCT-1999 9908-0157753.
PR 06-OCT-1999 9908-0157865.
PR 07-OCT-1999 9908-0158029.
PR 08-OCT-1999 9908-0158232.
PR 12-OCT-1999 9908-0158369.
PR 13-OCT-1999 9908-0159293.
PR 13-OCT-1999 9908-0159294.
PR 13-OCT-1999 9908-0159295.
PR 14-OCT-1999 9908-0159329.
PR 14-OCT-1999 9908-0159330.
PR 14-OCT-1999 9908-0159331.
PR 14-OCT-1999 9908-0159637.
PR 14-OCT-1999 9908-0159638.
PR 18-OCT-1999 9908-0159584.
PR 21-OCT-1999 9908-0160741.
PR 21-OCT-1999 9908-0160767.
PR 21-OCT-1999 9908-0160768.
PR 21-OCT-1999 9908-0160770.
PR 21-OCT-1999 9908-0160814.
PR 21-OCT-1999 9908-0160815.
PR 22-OCT-1999 9908-0160980.
PR 22-OCT-1999 9908-0160981.
PR 22-OCT-1999 9908-0160989.
PR 25-OCT-1999 9908-0161404.
PR 25-OCT-1999 9908-0161405.
PR 25-OCT-1999 9908-0161406.
PR 26-OCT-1999 9908-0161360.
PR 26-OCT-1999 9908-0161361.
PR 28-OCT-1999 9908-0161920.
PR 28-OCT-1999 9908-0161902.
PR 28-OCT-1999 9908-0161903.
PR 28-OCT-1999 9908-0162142.

Query Match 8.4% Score 286.5; DB 21; Length 504;
Best Local Similarity 25.2%; Pred. No. 3,56-18;
Matches 124; Conservative 63; Mismatches 202; Indels 103; Gaps 23;

QY 208 EPSTGIVSDRIYVTKIDMIVGIGLSMIVNGLIVVTC NIWFTSLYSSSDS-WIRGP 264
DB 53 dvnlgakpqlgls...dvnssqg;vnaflvstgtrjpaanaay;ssvovkwoyff 109
QY 265 PNCVVA-PGYSSATMSPQVPTTGSWSGVFEKNGEYSPSKTWTSLPAKVMPLTA 323
DB 110 qalaukrwytetqalipakliivigtr-----daloyeiipeagunckli 154
QY 324 DKQCLYKSD-----NHAMIFCW-----KKGSV-----PQAGYSTAMW 356
DB 155 dslllrqddpceunilyfwltdqellflannrsllspktyg;ko;pqjpeaanr 213
QY 357 YNTSGSDVKSACRPGSNPVPAPDAMGNAMVYDAVKCKLTFPSSPPOYSDATVNAHI 416
DB 214 ypaqssailipqilykpkkip.....ccg;vna;ssk-lady;ka;akki 257
QY 417 ITLGEIP-----TSPNTVFASNGLYFARTVHTSVVLWGSTFTRQDRGIP--FE 465
DB 258 y---cpaldkaf;rlinsakprkrltemf;primsdvl;lpnaad;lllnaukrwssayq 314
QY 466 DSHVFTPEIYVPEQ---LIFKGNPNSLIVVYHS;SLLLHGGVFRKAGG;GCMGTTN 521
DB 315 gkdpd;f;lllykbaatr;kr;tr;gl;kl;f;lf;v;ssail;lp;ek;v;vqasv;vdy;ky;v;v 374

PK	06-AUG-1999	990S-0147416.
PK	07-AUG-1999	990S-0147453.
PK	09-AUG-1999	990S-0147935.
PK	10-AUG-1999	990S-0148171.
PK	11-AUG-1999	990S-0148319.
PK	12-AUG-1999	990S-0148341.
PK	13-AUG-1999	990S-0148565.
PK	14-AUG-1999	990S-0148684.
PK	15-AUG-1999	990S-0149368.
PK	17-AUG-1999	990S-0149175.
PK	18-AUG-1999	990S-0149426.
PK	20-AUG-1999	990S-0149723.
PK	20-AUG-1999	990S-0149929.
PK	23-AUG-1999	990S-0149902.
PK	24-AUG-1999	990S-0149930.
PK	25-AUG-1999	990S-0150566.
PK	26-AUG-1999	990S-0150884.
PK	27-AUG-1999	990S-0151065.
PK	27-AUG-1999	990S-0151066.
PK	27-AUG-1999	990S-0151080.
PK	30-AUG-1999	990S-0151303.
PK	31-AUG-1999	990S-0151438.
PK	01-SEP-1999	990S-0151930.
PK	07-SEP-1999	990S-0152363.
PK	11-SEP-1999	990S-0153070.
PK	13-SEP-1999	990S-0153378.
PK	15-SEP-1999	990S-0154018.
PK	16-SEP-1999	990S-0154019.
PK	20-SEP-1999	990S-0154779.
PK	22-SEP-1999	990S-0155139.
PK	23-SEP-1999	990S-0155486.
PK	24-SEP-1999	990S-0155659.
PK	28-SEP-1999	990S-0156458.
PK	29-SEP-1999	990S-0156596.
PK	04-OCT-1999	990S-0157117.
PK	05-OCT-1999	990S-0157753.
PK	06-OCT-1999	990S-0157865.
PK	07-OCT-1999	990S-0158029.
PK	08-OCT-1999	990S-0158369.
PK	12-OCT-1999	990S-0159294.
PK	13-OCT-1999	990S-0159294.
PK	14-OCT-1999	990S-0159294.
PK	14-OCT-1999	990S-0159329.
PK	14-OCT-1999	990S-0159340.
PK	14-OCT-1999	990S-0159341.
PK	14-OCT-1999	990S-0159637.
PK	14-OCT-1999	990S-0159638.
PK	18-OCT-1999	990S-0159584.
PK	21-OCT-1999	990S-0160741.
PK	21-OCT-1999	990S-0160767.
PK	21-OCT-1999	990S-0160768.
PK	21-OCT-1999	990S-0160770.
PK	21-OCT-1999	990S-0160814.
PK	21-OCT-1999	990S-0160815.
PK	22-OCT-1999	990S-0160980.
PK	22-OCT-1999	990S-0160981.
PK	22-OCT-1999	990S-0160989.
PK	25-OCT-1999	990S-0161404.
PK	25-OCT-1999	990S-0161405.
PK	25-OCT-1999	990S-0161406.
PK	25-OCT-1999	990S-0161359.
PK	26-OCT-1999	990S-0161360.
PK	26-OCT-1999	990S-0161361.
PK	28-OCT-1999	990S-0161920.
PK	28-OCT-1999	990S-0161922.
PK	28-OCT-1999	990S-0161993.
PK	29-OCT-1999	990S-0162142.

QY	439	RTFHTSVMLPGSGTITFGAGRPDIP--FPDSTPVPPTPEIYVFPQ---DTFYKQNPSTIVR 433
DB	5	timdsdvtllpmdilllvadakrgcsjwtygkdpdptqpllykpkpazqvtrfqlkptlplr 64
QY	494	VHNSISLLLPVPRVFNKAGGAGTCTTN-HFDAGI----FTPWYVNSGGLNATFRKTR 548
DB	05	myhsaalllpdkkvzvgssatldyjkynccfpvctvckkspfydpalaa--rfkvt 122
QY	549	TST-QSVKVGKGRITISTD-----SSISKASLLRYCATHTVNTDQRIPLTLTN--NG 598
DB	123	larpkqkqyqltlucvda-kkqgqatdalkvmdagatcthszsmmrlllrrnqkpa 182
QY	599	GNSYSPQ--VPSDSGVALPQYMWLEVMNSAGVPSVASTIRV 637
DB	183	gqydlqgavqppqgslppyytlltdljk qpslscw-qf 222
RESULT 8		
AAG07598		
ID	AAG07598	standard; Protein; 224 AA.
XX	AAG07598;	
AC		
XX		
DT	17-OCT-2000	(first entry)
XX		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 4813.	
XX		
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25 FEB 2000; 2000EP-0201439	
XX		
PR	25-FEB-1999;	990S-0121825.
PR	05-MAR-1999;	990S-0121800.
PR	09-MAR-1999;	990S-0123548.
PR	23-MAR-1999;	990S-0125788.
PR	25-MAR-1999;	990S-0126264.
PR	29-MAR-1999;	990S-0126785.
PR	01-APR-1999;	990S-0127462.
PR	06-APR-1999;	990S-0128234.
PR	08-APR-1999;	990S-0128714.
PR	15-APR-1999;	990S-0129845.
PR	19-APR-1999;	990S-0130077.
PR	21-APR-1999;	990S-0130449.
PR	23-APR-1999;	990S-0130510.
PR	23-APR-1999;	990S-0130891.
PR	28-APR-1999;	990S-0131449.
PR	30-APR-1999;	990S-0132048.
PR	30-APR-1999;	990S-0132407.
PR	04-MAY-1999;	990S-0132484.
PR	05-MAY-1999;	990S-0132485.
PR	06-MAY-1999;	990S-0132486.
PR	06-MAY-1999;	990S-0132487.
PR	07-MAY-1999;	990S-0132863.
PR	11-MAY-1999;	990S-0134218.
PR	14-MAY-1999;	990S-0134219.
PR	14-MAY-1999;	990S-0134221.
PR	14-MAY-1999;	990S-0134370.
PR	18-MAY-1999;	990S-0134768.
PR	19-MAY-1999;	990S-0134941.
PR	20-MAY-1999;	990S-0135124.
PR	21-MAY-1999;	990S-0135353.
PR	24-MAY-1999;	990S-0135629.
PR	25-MAY-1999;	990S-0136021.

27-MAY-1999; 9905-0136332.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 01-JUN-1999; 9905-0137228.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139019.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139700.
PR 18-JUN-1999; 9905-0139769.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140359.
PR 23-JUN-1999; 9905-0140359.
PR 23-JUN-1999; 9905-0140364.
PR 24-JUN-1999; 9905-0140365.
PR 28-JUN-1999; 9905-0140827.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0141844.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142340.
PR 08-JUL-1999; 9905-0142807.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143621.
PR 15-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144087.
PR 19-JUL-1999; 9905-0144132.
PR 19-JUL-1999; 9905-0144133.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 23-JUL-1999; 9905-0144834.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145085.
PR 21-JUL-1999; 9905-0145089.
PR 21-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145216.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145274.
PR 27-JUL-1999; 9905-0145513.
PR 27-JUL-1999; 9905-0145513.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146386.
PR 03-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0146708.

PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148688.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151308.
PR 01-SEP-1999; 9905-0151940.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156556.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160817.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160880.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161459.
PR 26-OCT-1999; 9905-0161460.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 29-OCT-1999; 9905-0161993.

PP 19-MAY-1999; 99TTS-0134941

PR	21-MAY-1999;	99JIS-0135353
PR	24-MAY-1999;	99JIS-0135629

PR	25-MAY-1999	99JIS-11136021
PR	27-MAY-1999	99JIS-0136392

PR	28-MAY-1999;	9905-0130702
PR	01-JUN-1999,	9905 0137222

04-JUN-1999; 99US-0137502

08-JUN-1999; 09JTS-0138094

PR	10-JUN-1999;	9905-0138847
PR	14-JUN-1999;	9905-0139119

16-JUN-1999; 99US-01394533

PK 17-JUN-1999; 9905-0139454
PK 18-JUN-1999; 9905-0139454

18-JUN-1999; 99US-0139456

PR 18-JUN-1999; 99US-01394589
PR 18-JUN-1999; 99US-01394589

PR	18-JUN-1999	99US-0139460
PR	18-JUN-1999	99US-0139461

PR 18-JUN-1999; 99US-0139463

18-JUN-1999; 99US-0139763

PR 22-JUN-1999; 99JUS-0139899

PR	23-JUN-1999;	4911S-0140354
PR	24-JUN-1999;	4911S-0140695

PR	28-JUN-1999;	99US-0140991
PR	29-JUN-1999;	99US-0140991

PR 01-JUL-1999; 99US-0141842

02-JUL-1999; 99US-0142055

PR	08-JUL-1999;	99US-014280
PR	09-THU-1999;	99US-014292

PR 12-JUL-1999; 99US-01429/
PR 13-JUL-1999; 99US-0143542/

PR	14 JUL 1999;	2200 014400
PR	15-JUL-1999;	4415-014400

PP. 16-III-1944; 4415-114418

PR 19-JUL-1999; 99US-014433

PR	19-JUL-1999;	99US-014433
PR	19-JUL-1999;	99US-014433.

PR	19-JUL-1999;	99US-014432
PR	20-JUL-1999;	99US-014435

PR 20-JUL-1999; 99US-014488

PR 21-JUL-1999; 99US-014508

PR	22-JUL-1999;	99US-014508
PR	22-JUL-1999;	99US-014508

PR	22-JUL-1999;	99JIS-014508
PR	22-JUL-1999;	99JIS-014519

PR	23-JUL-1999;	99US-014514
PR	23-JUL-1999;	99US-014521

PR	23-JUL-1999;	99US-014222
PR	26-JUL-1999;	99US-014527

PR 27-JUL-1999; 99US-014591

100

PR	26-OCT-1999;	990S-0161360.
PR	26-OCT-1999;	990S-0161361.
PR	28-OCT-1999;	990S-0161420.
PR	28-OCT-1999;	990S-0161442.
PR	28-OCT-1999;	990S-0161993.
PR	29-OCT-1999;	990S-0162142.
Query Match		
Best Local Similarity	6.4%; Score 218.5; DB 21;	
Mismatches	64; Conservative 37; Mismatches 95; Indels 25; Gaps 10;	
CY	439 PTPPSSVTPMGSGFTFGGGRGPSP--FTHSTVFTRHLYVPPW--LIIFYKUNPSIYR 493	
DB	92 RIMSDTVLIPGGDILLIVGNAGKQSGWGYYKKAPAFAPILLYKPAARGLFRQLKPTLP 151	
OY	494 VHSLSLLPGCPVPFNFGGGLCGDCTTN-FEDAKI-----PTPNVLNSDGNLATRKIR 548	
DB	152 WYHSAAILLPQKXLYGSNTLDYKYKNGFFRLTGFSPYLDFIANI--RFKVI 204	
OY	549 TST-QSVKVGSRITISTD-----SSISKASLIRGTATHTVNTQGRPLRTIN--NG 598	
DB	210 LGLPKYKYGQILLVKVKDLKEKAUKNLIKVTMLPATHTSHISMETALLIYNVWPKA 273	
OY	599 CNSYSFO--VPDSGVALLPGYWMLFVNMSAGVPSVSTRV 637	
DB	270 GAGYALGAVPPNPGISPPYYIIIFAIYK-GVPSLGWIQY 309	
RESULT 10		
AAY35162		
ID	AAY35162 standard; Protein; 1146 AA.	
XX		
AC	AAY35162:	
XX		
DT	13-SEP-1999 (first entry)	
DE	Chlamydia pneumoniae cellular envelope protein.	
XX		
KM	Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;	
KW	sinusitis; purulent otitis media; erythema nodosum; pharyngitis;	
XX	vaccine; neutralising epitope.	
OS	Chlamydia pneumoniae.	
FN	W09927105-A2.	
PD	03-JUN-1999.	
XX		
PF	20-NOV-1998; 98MO-1B01890.	
PR	04-NOV-1998; 98US-0107078.	
PR	21-NOV-1997; 97FP-0014673.	
PA	(GENSET) GENSET.	
PI	Grlfais R;	
PR	WPI: 1999-357842/30.	
XX		
PT	Genome sequence of Chlamydia pneumoniae	
XX		
PS	Page 1627-1629; disclosure: 1912pp; English.	
XX		
CC	AAY4584-Y35879 represent the proteins encoded by all the open reading	
CC	frames in the complete genome (see AAX41990) of Chlamydia pneumoniae.	
CC	C. pneumoniae causes respiratory disease such as pneumonia and	
CC	bronchitis and is thought to be a contributing factor in heart	
CC	disease, sarcoidosis, sinusitis, purulent otitis media, erythema	
CC	nodosum or pharyngitis. The polypeptides encoded by the open reading	
CC	frames of the C. pneumoniae genome (see AAY34584 Y35879) can be used in	
CC	immunogenic compositions as vaccines. Vectors containing C. pneumoniae	
CC	nucleotide sequences can also be used as immunogenic compositions,	
CC	especially where the vector directs the expression of a neutralising	

530 tpehmgclvstetlrlmkllr...trgpydytrssmjffdljgocfgy 548
 457 GGRRCPEPEEPIEFTT...EIVYELDTFKQKNSIVPYHSLIPQPV 567
 589 -----tdggnitlplvstlryy...-----gdlitrggqf...-dgvf 627
 508 FNDGDMQDNTHHEFAEELINVTYNSQNMATPPTTFPTGQVAVGPTTSTSS 567
 628 y---qtqfcllssagcllrlilngslgtatqlvqfssclsa---volasqvl 680
 505 LSAASLLKALIAHNTLCHLSTTTH...TGVYSGVAFGFSVATFG 614
 681 lnaqgvstf---chlttrmgdqlplstvgstvsptgpaqmvllnaqglcltg 735
 617 y 617
 6b 736 y 736
 RESULT 12
 AAB93143
 ID AAB93143 standard. Protein; 521 AA.
 AC AAG75453;
 DI 03-SEP-2001 (first entry)
 DE Human colon cancer antigen protein SEQ ID NO:6217.
 FE Human colon cancer colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma. Chromosome 1.
 LS Homo sapiens.
 OS W2203123920-A2.
 PR 05-APR-2001.
 PP 28-SEP-2000; 2000MD-030524.
 PR 26-SEP-1999; 99US-015717.
 PR 03-NOV-1999; 99US-016326.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PL Ruben SM, Barash SC, Birse CE, Rosen CA;
 OR W21; 2001-23537/24.
 DR N-PSDB; AAB34856.
 XX Nucleic acids encoding 1377 human colon cancer-associated polypeptides,
 useful for preventing, diagnosing and/or treating colorectal cancers -
 Claim 11; Page 7661-7664; 9803pp; English.
 XX AAB93143 to AAB93145 and AAG75453 to AAG75455 represent human colon
 cancer-associated nucleic acid molecules (N) and proteins (P), where
 the proteins are collectively known as colon cancer antigens. The colon
 cancer antigens have cytosolic activity and can be used in gene
 therapy and vaccine production. N and P may be used in the prevention,
 diagnosis and treatment of diseases associated with hyperplastic P
 expression. For example, N and P may be used to treat disorders
 associated with decreased expression by facilitating mutations or deletions
 in a patient's genome that affect the activity of P by expressing
 inactive proteins or to supplement the patient's own production of P.
 Additionally, N may be used to produce the colon cancer-associated P,
 by inserting the nucleic acids into a host cell and culturing the cell
 to express the proteins. N and P can be used in the prevention, diagnosis
 and treatment of colorectal carcinomas and cancers. AAB93143 to AAB93145
 and AAG75453 represent sequences used in the exemplification of the
 present invention.
 CC N.B. Pages 666 to 682 and page 7653 of the sequence listing were
 missing at time of publication, meaning no sequences are present for

66 SEQ ID NO 1027 to 1052, 7921 and 7922
 XX
 SQ Sequence 521 AA;
 Query Match 4.2%; Score 142.5; DB 22; Length 521;
 Best Local Similarity 19.1%, Prev No. 0.022;
 Matches 93, Conservative 69, Mismatches 170, Indels 155; Gaps 22;
 28 NKAIDGNKDTFHHFFYGANGDPKPHPTTYTIDKTTQNVNGLSMLPRQDQNGWIGRHEV 87
 108 kklldgallldgaefgsddd...higfvqkpp-----premhkq----- 145
 88 YLSSDGTWGSIVAGSCMPADSTTYSNFETPPARYVRLVATFENQSPWISIAE----- 143
 146 -lsssst-----gclsspsatvgs...pkhwkivasektsnltylelavldqif 191
 144 -----NVFOASSYTAPOPLG-----RWGPTTDLPIV-----AAALPPTSGRVIM 185
 192 cvlthgrnspgspslspklskslscfmgjdellekcpmpgyarsgltaemgklla 251
 146 WSSYVNOAFASISPSGTTILTSMPSTGIVSCRTVVTIKRMFTPSISKGRGQIVTTCN 245
 252 aggyrocc-----lrvocyphldhwsflapmtrparfmgavlm--gqlyvrggs 302
 246 LAKKISL-----YSSSSSLMPLGPMQVAFYSSSAIMWSDCVHILDSWS--GVYERKG 299
 363 tghshllsvgeayhshldhldwlpvpeflrncagvcdlngkllyvgsgdpygqgylknc 361
 300 EYVSPSSKRTWTSILFNKV-----NPLTADKQSIYKSNHAW 336
 362 dvlldpvlkiklwsclpnlrlthgsavcejgqlyllggaesncdlrvc---ynpennlw 418
 337 LFGWKKGSVFQAGSPSTAMWVYVSSGIVKSAKKQSNVGVAPDMWGN-----AV 387
 419 cl-----iapm-varrgqvavlngkl-----fvcgtdgshalscve 456
 388 MYDAVAKGKILIFGSPDYDDSDATINAHILICEPGTSNPTVFSANGLYFARTFHTSVL 447
 457 mydplrnekwkmgm-----tsprnagiatvg-----nltyavvgf----- 493
 448 PGGSTFI 454
 494 -dgnelf 499
 RESULT 13
 AAB93143
 ID AAB93143 standard. Protein; 600 AA.
 AC AAB93143;
 DI 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:12047.
 FE Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KW Homo sapiens.
 OS EP1074617-A2.
 PR 07-FEB-2001.
 PP 28-JUL-2000; 2000EP-0114126
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241849
 PA (HELI-) HELIX RES INST.

1 ASAPTGSATSPNNWAVTQDSAQSGNECNKAIDGNKDTFWHTFYCANGDPKH

227 331VWGG NDAA...
 126 01VVG00NNDAA...
 283 3NSDVFEN...
 185 36FENEL...
 339 66H0A...
 211 6YK...
 392 3VKR...
 284 3V...
 449 3...
 441 3...
 507 3...
 401 3...
 560 3...
 456 3...
 607 3...
 516 3...

RESULT 4

15935

probable galactose oxidase (EC 1.1.3.9) (F5620.250 [similarity]) - Arabidopsis thaliana
 CSpecies: Arabidopsis thaliana (mouse-ear cross)
 CDate: 04 Feb 2000 #sequence: 141602000 #cds: 141602000 #cds: 141602000
 CAccession: A145993
 CMonot: A145993
 AReference number: 22017
 AAccession: 145935
 AStatus: preliminary
 AIsolate type: DNA
 AIsolates: 1:545 -HBN
 AReferences: EMBL:AL142660
 AExperimental source: cultured cell line: HAC clone F5620
 CGenetics:
 AMap position: 3
 AProbe: F5620.250
 CSuperfamily: Arabidopsis thaliana hypothetical protein F5620.250
 CKeywords: copper metal binding, oxidoreductase
 CFile: 145935 Link: cytology: 145935 #status: predicted
 CFile: 145935 Link: cytology: 145935 #status: predicted

Query Match 7.8% Score 2661 DB 2: Length 545
 Best Local Similarity 27.0% Pred. No. 2: 06-07
 Matches 126 Conservative 67 Mismatches 184 Indels 90 Gaps 28
 227 331VWGG NDAA...
 126 01VVG00NNDAA...
 283 3NSDVFEN...
 185 36FENEL...
 339 66H0A...
 211 6YK...
 392 3VKR...
 284 3V...
 449 3...
 441 3...
 507 3...
 401 3...
 560 3...
 456 3...
 607 3...
 516 3...

385 3...
 274 3...
 438 3...
 328 3...
 493 3...
 388 3...
 441 3...
 595 3...
 498 3...

RESULT 5

H6278

hypothetical protein AAF43943.1 [imported] - Arabidopsis thaliana
 CSpecies: Arabidopsis thaliana (mouse-ear cross)
 CDate: 02 Mar 2001 #sequence: 141602000 #cds: 141602000
 CAccession: H6278
 CMonot: H6278
 AReference number: 22017
 AAccession: 145935
 AStatus: preliminary
 AIsolate type: DNA
 AIsolates: 1:545 -HBN
 AReferences: EMBL:AL142660
 AExperimental source: cultured cell line: HAC clone F5620
 CGenetics:
 AMap position: 1
 AProbe: F5620.250
 CSuperfamily: Arabidopsis thaliana hypothetical protein F5620.250

Query Match 7.2% Score 245.5 DB 2: Length 564
 Best Local Similarity 23.5% Pred. No. 5: 7e-08
 Matches 128 Conservative 79 Mismatches 221 Indels 117 Gaps 27
 197 3...
 33 3...
 222 3...
 93 3...
 254 3...
 153 3...
 309 3...
 213 3...
 367 3...
 267 3...
 422 3...

[illegible][illegible]

RESULT 13
13443
Hypothetical protein K06A9_1a - *Caenorhabditis elegans*
Cloned: 29-05-1996 #E16026-10-0126 29-05-1996 #E16026-10-16-1996
ClAccession: 134434
RefSeq: C2_Gallunay, S.
Submitted to the EMBL data library, December 1996
Description: The sequence of C. elegans cosmid K06A9.
Reference number: 221525
Accession: 134434
Access: 13-01-1997 10:00:00
Molecule type: DNA
A residues: 1-2212 CDS
A CDS reference: EMBL 200446 PIRN A0700011 OSTR 0200028 CESP K06A9_1a
A EMBL entry source: strain: Bristol N2 clone: K06A9
ClGenetic:
A name: CESP:K06A9_1a
A map position: X 2753 1000 1000 1000 1000 1000 1000 1000 1000 1000
A features: 13443 1000 1000 1000 1000 1000 1000 1000 1000 1000

```

RESULT 14
T00349
A:Accession: T1: Aspergillus aculeatus
C:Species: Aspergillus aculeatus
C:Date: 01-Feb-1999 #sequence-revision: 01-Feb-1999 #text-change: 16-Jul-1999
C:Accession: T00349
C:Author: M.; Takada, G.; Kawaguchi, T.; Sumitani, J.
Submitted to the EMBL Data Library, June 1998
A:Description: Avicelase III from Aspergillus aculeatus.
A:Feature: number: 31441
A:Accession: T00349
A:Status: Preliminary; translated from GR/EMBL/DBJ
A:Modification: type: mRNA
A:Protein: 1956, ARA
A:Cross-references: EMBL:AB015511, NID:J193887; PID:d1029971
C:Genetics:
A:Gene: avicIII
A:Suppl:frag: fungal cellulose binding domain homolog,
F_823 8-4/Dominant fungal cellulose-binding domain homolog, (FCB)

```

Query Match 4.7% Score 159.57 DB 2: length 856.
Best Local Similarity 20.0% Pred. No. 31924.

Matches 149: Conserved 75: Mismatches 299: Indels 19: Gaps 32:

```

41 TPVCAALFPHHLYTLPETLAVNRLSMLEPCNENQWZPHVYSSQ QNW 96
112 IYVYGG-----LNDLV-----VIGLIMV AALNMPDRLKSDSGALWSL 356
57 GSF-----VASGWEALSTIKYSNEELIPARYVLAITANZPWTSAEI 143
557 WENQYESTINVYSYDLSNAWICL-TSTIGLTPRVNMMWALALPEPSNHWLYGTL 415
144 NVDGSSYVIANVGLCWGPTTGLFIVAAALFPLGCRVLMWSYRNDAPGSGCITL 203
416 IYVGHML-----NWSKNHNVKSLAVGLEPMVLCGL-----IPGHAL 458
204 TSSMDPSTGLVSRVIVIKHMDP-----GLSMQEN--GQIVVTGQNAKKT-SLYD 254
459 ISAVGCGGGRHSHLAAFNAYHPIYGLINQIYVGNKPSNIVHSSAGVYPIALSS 518
255 EAGSKTGLMLVAFVTSNMTT TETTLNNGGVV FGLFVNSKRWLSLH 314
519 NQSTWATPAASTISIGAVASAMGVH MESTSALVSK-----SGTITAVSS 571
415 AFVPMELAEGLIYPSINHWI PWWFVSF-----GAGSLAMWY--V1 464
572 LSTAVIASK-----SINVF-VYSAGAIYVGNATSTFTYSGSTTVNAPRHP 525
469 SGSDVSKAKH-----GSKGVAVIOMV-----NAVMYAVVOKH1E03 401
626 SIADVASTIKGIMSHSTVOSTPTQOSVTAQSPKRASTGSVVYGEFTIDCA 685
402 STVYGSTATNAHILDELSTSNIVASNGVIARLH SVLHPSIT1332EFG 461
606 AGPSEELAGINWVILASGFGSGANVNG-----DLQTY--GDFVPS 729
462 LPEESTFVTELYHEGLPYKGNNSIVVYVHSILHSGVFNQ333Q3Q11N 521
730 -----HERPHILKQ-----SQPREPVR--HDDN--GDTTTS 758
522 HFLAQIFTPNVIYNSKGNLAPRPLTPTSTSVVGVPTTSTSSISKASI1RGTATH 581
759 KISTVST-----TKITTSASTSSSTTV--KITTSSTTSKAST--TTTK 804
582 TANDGRKRLITLTH -----GSSYSILVETSTVA:PSYMLP--VMNSNG 627
805 LITIS-----TTSSGTLIASVAGCGNGW-----TGATVCTGTCTCYSNAPYSG 853
628 VPS 630
854 VPS 856

```

RESUME 15

130261
Title: (E: 3.2.1.14) homolog - Streptomyces lividans
Species: Streptomyces lividans
Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #extl_change 18-Feb-2000
Accession: 130261: 130261

Author: J. N. Eichensow, G. Vici, P. Plöndl, W. Altenbuchner, J.
Submitted to the EMBL Data Library, July 1998
Accession number: Z20798

Accession: 130261

Accession: preliminary: translated from cR/FMD/20BJ

Accession: preliminary: translated from cR/FMD/20BJ

Accession: preliminary: translated from cR/FMD/20BJ

Accession: preliminary: translated from cR/FMD/20BJ

Accession: preliminary: translated from cR/FMD/20BJ

Accession: preliminary: translated from cR/FMD/20BJ

Accession: preliminary: translated from cR/FMD/20BJ

C:Genetics:
A:Note, amplifiable element AUDI
Keywords: glycosylase, hydrolase

Query Match 4.7% Score 159.57 DB 2: length 1444
Best Local Similarity 21.0% Pred. No. 0.048;
Matches 130: Conserved 86: Mismatches 279: Indels 147: Gaps 33:

```

15 VPDASQNECKKATGKDEIFWHFYANDEKPHYITLTKTQNNVSLM-LRQ 74
187 IASSLQAVLPENNGNAPYAFSANN-----PQWIDAPGSAAPRVAVI PLD-- 249
75 IENQRMNCHH--VITSEGTNNSTFVANGSGWFASTKYS---NHTPRVAVHVA 128
240 ----LQMPASQILKIQASDNGSDPDLITAAKATTPVAGGQSAVTEPDATVAVYR-VL 294
126 TPANCPWTSAIINVEFGSSTAPQDGLPMDPTLDDPVPAAAL-EPTSCR-VLM 186
295 VIANDLPAQISLEY-----GPTD-----GAGVADANLPVADIDGLH 342
187 SSTRQATGGLSGLLISW4ELGVSC4L VITD46 113E-2K3 257
343 NSSDGL-----AVIGVYANMELI SVAGVITVYDIPMAGITVYVPRADAMNGSD 399
236 QIVVQGNAPVISTYSSS-----SWLGGIMVAVGVGSATKGRV- 281
400 SSVTPPATGTCATPASNALTEVPQVPLTWASTPDEAVTGD---VYALNPL 456
294 PTTGQWSGVFEKQDYS--PSEPTWETPRVAVHMTAVVGVYVST PMWFP 439
457 KVAAG-----NVTYDIDQPASANVTFVPAK-----IAAGNSDSNSVFG 500
44 WFTSVGLAAGFSA---MMVYTSAGVAFSA:RHS9:VAVGAM GGNVMTAV 552
591 TRGNSAVKPIIASSHIFVAFNKNQSTIYVAGANQVNI1111CANAVNSV 560
493 KCHLLEGGSPDYUSLDTNAHITLGEPTSNIVFASNGLYF-APHTPSVVI1DQ- 459
561 VKL-----NPDAGWSPRTGHEVLGPAQASGLTSLAAKDVVFDPASNSV1PVG 615
451 -----SPTTGGGPPGTPE-DSITVTPPPIYVPRQPPFYGNNSIVRVVHST 498
616 RAADVQIKYTSNGATAGQ--VAPQVIGTPAPMDLEV-----GLTTPSGLPV----ES 665
499 ALIFGFEVNDGGGLGCTINHECAITFNITNSGMLATEK1TRTSTSV---K 555
666 DRLTVAIVNNSGA-----DAVA-----GVAVRKADIVATADIVALE 705
556 VGGRIITSDSSISKASLLRYGTATHVNDQPRHPLILNNGNSYS 603
706 AGMORTVASLGAAREAGSELSAVADEAN-----ELIENETNNVT 747

```

Search completed: January 3, 2002, 10:02:27
Job time: 122 sec

GenCode version 4.5
Copyright (c) 1997 2000 Comagen Ltd.

Gen protein - protein search using SW model

Run on: January 3, 2002, 10:01:30 : Search time 13.31 seconds
(without alignments)
1756.281 Million cell failures/sec

Title: US-09-722-602-10
Perfect score: 3425
Sequence: 1 ASAP10SALFORMNMAVTFDSE.....LVVNSGVHVSASTLRVQ 639

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 100059 seqs, 36664827 residues
Total number of hits satisfying chosen parameters: 100059

Minimum (8 seq length) 0
Maximum (8 seq length) 200000000
Maximum DB seq length 1000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the local score distribution.

SUMMARIES

Result	Score	Query No.	Match	Length	FR	ID	Description
1	3425	94.9	680	1	GAOA_TASTE	Q01745	dactylium t
2	293	6.8	717	1	NCYM_MTV	A09934	anti-mouse
3	209	6.1	1014	1	NAMH_MOUSE	P29767	clathrin
4	149	4.4	1723	1	Q92812	chlamydia p	
5	142.5	4.2	1902	1	FCF_IATRA	Q92470	chlamydia p
6	141.5	4.1	751	1	KHLL_MOUSE	Q93174	mus musculu
7	141.5	4.1	969	1	NID2_YEAS	Q93629	saccharomy
8	140.5	4.1	823	1	XAMH_MANE	Q03106	xanthomonas
9	138.5	4.0	748	1	PAH1_HUMAN	Q09809	podospira a
10	127	4.0	1256	1	HEFI_HUMAN	Q06653	r outer mem
11	137	4.0	1656	1	OMPH_PICJA	Q03155	eschlerichia
12	136	4.0	1295	1	ALVA_PVLI	Q25485	limulus pol
13	135.5	4.0	915	1	SOKH_LIMPO	P87061	schistosom
14	132	3.9	1117	1	TEAL_SCHPO	P33666	schistosom
15	132	3.9	2097	1	TEAL_SCHPO	P33666	schistosom
16	131.5	3.8	1902	1	NCYM_MTV	Q09809	podospira a
17	131.5	3.8	2090	1	N214_HUMAN	P35658	homo sapien
18	130	3.8	1016	1	NYAT_YEAS	Q03106	xanthomonas
19	125.5	3.6	1333	1	AD41_PVLI	P33160	eschlerichia
20	129	3.8	646	1	PAH1_HUMAN	P08675	eschlerichia
21	127	3.8	1232	1	TEAL_SCHPO	P33666	schistosom
22	126.5	3.8	926	1	TEAL_SCHPO	P33666	schistosom
23	126.5	3.8	1902	1	NCYM_MTV	P35658	homo sapien
24	126	3.7	619	1	LAO1_NEUTR	P06911	neutrospora
25	126.5	3.7	1090	1	Q9X2_CELFI	P06911	neutrospora
26	126.5	3.7	1090	1	Q9X2_CELFI	P06911	neutrospora
27	126	3.7	722	1	Q9X2_CELFI	P06911	neutrospora
28	125.5	3.7	722	1	Q9X2_CELFI	P06911	neutrospora
29	125.5	3.7	722	1	Q9X2_CELFI	P06911	neutrospora
30	125.5	3.7	722	1	Q9X2_CELFI	P06911	neutrospora
31	125	3.6	920	1	N214_HUMAN	P35658	homo sapien
32	125	3.6	1396	1	AD41_PVLI	P33160	eschlerichia
33	124.5	3.6	659	1	PAH1_HUMAN	P08675	eschlerichia

34	124.5	3.6	1400	1	120K_PICPI	P14914	homo sapien
35	124.5	3.6	1337	1	PTPQ_HUMAN	Q12913	homo sapien
36	124.5	3.6	1654	1	OMPH_PICPI	Q03047	r outer mem
37	123.5	3.6	1052	1	MGPC_MYCEP	P03747	mycoplasma
38	123	3.6	1754	1	PMPB_CHLTR	Q84418	chlamydia t
39	121.5	3.5	1569	1	YPIA_ECOLI	P52143	eschlerichia
40	121.5	3.5	1802	1	HKR1_YEAST	P41809	saccharomy
41	121	3.5	426	1	V54_RPM13	P01668	bacterioph
42	121	3.5	1276	1	PMP6_CHLPI	Q92894	chlamydia p
43	120.5	3.5	918	1	SCPA_LIMPO	Q25390	limulus pol
44	120.5	3.5	1016	1	EMPH_CHLTR	Q84860	chlamydia t
45	120.5	3.5	1672	1	PMH1_CHLMO	Q94792	chlamydia m

ALIGNMENTS

RESULT 1
GAOA_LINCDE
ID GAOA_LINCDE STANDARD: PRT: 680 AA.
AC Q01745;
DT 01-jul-1993 (rel 26, last sequence update)
PT 01-jul-1993 (rel 26, last sequence update)
DE 20-AUG-2001 (rel 40, last annotation update)
DE GALACTOSE OXIDASE PRECURSOR (PG 1.1.3.9) (GAO)
GN GAOA.
OS Dactylium dendroides (Cladobotryum dendroides).
OC Eukaryota, Fungi, Ascomycota, Pezizomycotina, Sordariomycetes;
OC Hypocerales, Hypocerales, Hypocerales.
OX NCBI_Taxid=5132;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
FX MPM178F-0235025; PubMed=156070;
FA McPherson R T., Ogel Z R., Stevens G F., Yadav K D S., Keen J N.,
FA Knowles P.F.;
PT "Galactose oxidase of Dactylium dendroides. Gene cloning and sequence
RT analysis."
PI J Biol Chem 267:8146-8152(1992).
FN [2]
FP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
FX MEDLINE=9163641; PubMed=2002850.
FA Lee N., Phillips S.E.V., Stevens G F., Ogel Z R., McPherson M J.,
FA Keen J N., Yadav K D S., Knowles P.F.;
PT "Novel thioether bond revealed by a 1.7 A crystal structure of
RT galactose oxidase."
RL Nature 350:87-90(1991).
CC -- FUNCTION: CATALYZES THE STEPS-SPECIFIC OXIDATION OF A BROAD RANGE
CC OF PRIMARY ALCOHOL SUBSTRATES.
CC -- CATALYTIC ACTIVITY: D-GALACTOSE + O(2) = D-GALACTO-HEXODIALDOSE +
CC H(2)O(2).
CC -- COFACTOR: COPPER.
CC -- SUBUNIT: MONOMER.
CC -- SUBCELLULAR LOCATION: SECRETED.
CC -- SIMILARITY: CONTAINS 3 KETCH REPEATS.
CC ! ENZYME: NAME: Maltinase enzyme manual,
CC WWW: http://www.washington-biochem.com/manual/5/GAO.html".
CC This SWISS-Prot entry is copyright. It is provided through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See: http://www.ebi.ac.uk/submit/)
CC or send an email to license@esb.sib.ch).
CC EMBL: M86819, AA016228.1; .
CC PDB: 1GOG, 3I-JAN-94.
CC EPR: 1GOG, 3I-JAN-94.
CC PDB: 1GOG, 3I-JAN-94.
CC InfoProc: TP000421, FA58_C.
CC InterPro: IPR001798; Ketch.
CC Pfam: PF01344, Ketch; 3
CC PRINTS: PR00501, KETCHREPEAT.

SMALL: SM0231: PASNO: 1.
X-ray structure: Strand: copper: Repeat: 3D-structure:
STAND: 1 41 GALACTOSE OXIDASE
STAND: 42 680
REPEAT 279 321
REPEAT 323 372
REPEAT 492 544
DISULFID 59 68
HELIX 413
METAL 413
METAL 413
METAL 536
METAL 537
METAL 622
METAL 48 49
STAND 56 59
STAND 63 63
STAND 64 65
STAND 66 66
HELIX 72 72
STAND 76 77
STAND 81 81
HELIX 85 87
STAND 94 112
STAND 120 121
STAND 123 123
STAND 125 131
STAND 142 145
STAND 153 159
STAND 172 172
STAND 174 174
STAND 181 181
STAND 183 188
STAND 197 198
STAND 201 207
STAND 214 217
STAND 219 221
STAND 224 227
STAND 236 237
STAND 243 249
STAND 250 253
STAND 254 260
STAND 262 263
STAND 271 274
STAND 276 277
STAND 279 283
STAND 285 286
STAND 288 289
STAND 291 295
STAND 296 299
STAND 300 303
STAND 312 312
STAND 315 318
STAND 320 321
STAND 324 327
STAND 330 330
STAND 342 342
STAND 344 344
STAND 345 348
STAND 349 353
STAND 354 355
HELIX 359 361
STAND 366 368
STAND 372 372
STAND 373 373
STAND 376 380
STAND 382 384
STAND 388 388
STAND 394 399
STAND 402 403
STAND 405 411
STAND 413 414
STAND 415 416

FT STRAND 417 418
FT STRAND 423 423
FT TURN 424 424
PT STRAND 426 431
PT TURN 432 435
FT STRAND 436 440
FT STRAND 443 443
FT STRAND 444 443
FT STRAND 450 450
FT STRAND 452 452
FT STRAND 455 459
FT TURN 463 464
FT STRAND 468 471
FT TURN 473 474
FT STRAND 474 481
FT STRAND 480 487
FT TURN 489 490
FT STRAND 493 496
FT STRAND 499 499
FT STRAND 501 501
FT TURN 504 505
FT STRAND 510 510
FT STRAND 515 518
FT TURN 519 522
FT STRAND 523 526
FT TURN 536 537
FT STRAND 548 542
FT TURN 544 545
FT STRAND 548 552
FT TURN 557 558
FT STRAND 565 570
FT HELIX 572 574
FT STRAND 575 575
FT TURN 577 578
FT STRAND 581 581
FT STRAND 587 590
FT STRAND 594 596
FT TURN 597 598
FT STRAND 600 605
FT STRAND 611 616
FT STRAND 619 621
FT TURN 622 623
FT STRAND 624 624
FT STRAND 630 633
FT STRAND 636 637
FT STRAND 642 646
FT TURN 651 653
FT STRAND 657 664
FT TURN 666 667
FT STRAND 670 670
FT STRAND 674 679
SQ SPODENCE 680 AA: 72823 MW: 2970561B63E46E9 CMC64:

Query Match 99 98: Score 3420; DR 1: Length 680;
Best Local Similarity 99.8%, Pval: 2.5e-205;
Matches 638; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 ASADIGSAISRNNAVTCDSAGSNGKKAIDCKKDIPIWHIFPYANGDPKPPHYITIMK 60
DB 42 ASADIGSAISRNNAVTCDSAGSNGKKAIDCKKDPWHIFPYANGDPKPPHYITIMK 101
OY 61 TTONVNLISMLPQDQGNONIGFHEVYLSGDTNWSPPVAGSGWFAUSTIKYSNFEIRP 120
DB 102 TTONVNLISMLPQDQGNONIGFHEVYLSGDTNWSPPVAGSGWFAUSTIKYSNFEIRP 161
OY 121 APYVLAITEANQPTSTIAEINVPQASSYIAPQPLGHWPTITLPIVPAANAIPYTS 180
DB 162 APYVLAITEANQPTSTIAEINVPQASSYIAPQPLGHWPTITLPIVPAANAIPYTS 221
OY 181 GPVLMSSYNDAGGSPGDTTITSSWDPSTGIVSDRTVIVIKHIMKPGISMGNDQIV 240
DB 222 GPVLMSSYNDAGGSPGDTTITSSWDPSTGIVSDRTVIVIKHIMKPGISMGNDQIV 281

[illegible]

D6	76	--PCCVAVMAGTGGNNQ--DMSDKOYATKIMDPVKOTLKK--VLPSPIDPCTGHTOLA	128
OY	246	NGLIVITVG-----	244
D6	129	NNNIIIAOSTRHYEKLADVAKNGLMVHMLIQRITTFACETTCDEIKETVSFR	188
OY	245	-----ND-----	246
D6	169	VLVPRKEVLEDAIALVKNLEIGERLYVEACKKGAYETETINTELOLSGNAMNY	248
OY	247	AARSTL-----YSSSSISWIPHPMGVARGYSANSDRPETIG-SW	290
D6	219	AAALAIPTVDICIDALEPEENVAKEYIKVRIMHEARKPTITTTGIKKILSVGLHD	308
OY	291	SQVYEKSEYESSPSKTIWTLSPNAKVPMILADRGLYSDINHAMELGKGSVFQAR	450
D6	309	IQQIIVGRNEVYDIKTKAMVITDRKQPPIYA-----LFLLMNCKIPYSG-	354
OY	351	SIAMMYTSSGNGVKSAGKRQSNRCVAPIAMCGMAVMYAVANGKILLPGSPDYQUSDA	410
D6	355	-----AANGQPDVVORTPGIWDVETNK---FLKVGMSDAM	389
OY	411	TTHAHITIGEPOTSNTVFASNC-----LYPAFTF	442
D6	390	LFTAFNVLLPFAQDKRYAVICGGVCYCNISKSEKPLADLKADPKFVDSLEKTRYP	449
OY	443	TSVVLIDGSTFTTGQR--KQIPEDSFPVTTPELYEQDTFKOMPNSIVRYXISIL	500
D6	450	OMSLIPDDSVLSCGSGSEYRG--KQSNLIQAANIETIRHHQVADIVCRBHSSTI	506
OY	501	LLEPQR-VNAGGGGLGD---CLINAE AQIELECYLRSDNLAIRKRTISTQSV	554
D6	507	LLEPKLMFGDSLALAKANKRKCFQLDLALILEFYLD-----SRPLS-GHQFI	560
OY	555	KWGKHITIST-ISSISKAASLIRGLAHVHTGRRRLTTLNRGRNRSF-YPSISGV	612
D6	561	AWGSHFTSPASTVSKRVKVELLPASASTHVTDQRSIALDFAD-GRKLIVPTGRNL	619
OY	613	ALPYGMMLPYNMASGVPSASTIRV	637
D6	620	VOSQWMEVTVDGCPISKAEVAV	644
RESULT	3		
OYFG4		PRELIMINARY:	PRI: 615 AA.
ID	OYFG4:		
PC	01-MAR-2001 (TREMBLrel. 16; Created)		
DT	01-MAR-2001 (TREMBLrel. 16; Last sequence update)		
DT	01-MAR-2001 (TREMBLrel. 16; Last annotation update)		
DE	FINZ1.11.		
OS	Arabidopsis thaliana (Mouse-ear cross).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.		
XX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Johanson-Hopson C., Brooks S., Buchler E., Chao G., Khan S., Kim U.,		
KA	Shim P., Altai H., Hei Q., Chin C., Chou J., Choi E., Conn L.,		
RA	Cenky A., Goncharov A., Hansen N., Howing R., Koo T., Lam B., Lee J.,		
RA	Leuz G., Lu J., Liu A., Liu K., Liu S., Mukharshy B., Nguyen M.,		
RA	Falm C., Pham P., Sakano H., Schwartz J., Southwick A., Thayer A.,		
RA	Torilmi M., Vaysberg M., Yu G., Fiederspiel N.A., Theologis A.,		
RT	Ecker J.R.;		
RT	"Genomic sequence for Arabidopsis thaliana HAC FINZ1 from chromosome		
EL	1";		
EL	Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.		
EL	[2]		
RP	SEQUENCE FROM N.A.		
RA	Ecker J.R.;		
EL	Submitted (MAY-1997) to the EMBL/Genbank/DDBJ databases.		


```

> Best local similarity: 24.5%, Pval: 5.3e-26, 95%
> Matches: 54, Conservation: 25, Mismatches: 69, Indels: 56, Gaps: 4
>
> 4 PIGASISNNNAVCTDAASC .....
> 1162 PAFAGTSATWVIAAVMEIATRTVSGAAAVTDEPLASTATADSENTSSGAE 1221
>
> 25 NECKAIDAKNDOTWHITTYGAANDKPRPHATYTEMKTTONVMSI PRGDNDQNGTGR 84
> 1222 TSVNVDITATITHTHIIIGCAVYVIMWTKIQTAAI VGGVYVYVGGV 1280
>
> 85 HEVYISSTGMSGCVAVGVAVATGTSRPTTFAPVYFVATFAPAGDPTGTAFIN 114
> 1281 YEVAVSDGDEAK-TTVAATGTLKRVETGVSSTGVASVSTFALLANQGVAAAEPR 1339
>
> 145 VFOASSYIAPOHGHMCHPTDIP 168
> 1340 VY ..... GVPVPLP 1348
>
>
> RESULT 12
> ID Q12M3 PRELIMINARY PRT 446 AA
> AC Q12M3
> DI 01-MAR-2001 (TEMBLEL 16, Created)
> DI 01-MAR-2001 (TEMBLEL 16, Last sequence update)
> DI 01-JUN-2001 (TEMBLEL 17, Last annotation update)
> DE HYDROLICLICAL PROTEIN PA1874.
> ON PA1874.
> OS Pseudomonas aeruginosa.
> OT Bacteria: Proteobacteria; gamma subdivision: Pseudomonadaceae;
> OC Pseudomonas.
> OX NCBI_taxid=287;
> RN 111
>
> SEQUENCE FROM N A
> K0 SEIRINE-PA01;
> EX METALINE-2043737; PubMed=1656444;
> KA STORCK C.K., PAUM X.-C.T., EYIN A.L., MIZOGUCHI S.D., WARRENER P.,
> HILGERT M.J., HILGERT M.F., HILGERT W., KOWALYK G.T., LAGROU M.,
> GARTER R.L., GARTER L., LAGROU P., MONTROCK-WATSON S., YOUNG Y.,
> ROTH J.L., GARTER F.N., FINE J.E., YOUNG A., LAGROU P., LAGROU M.,
> SMITH K.A., SPENCER E.H., WOOD G.K.S., MO Z., PAULSEN I.T.,
> BELZER J., SAYER M.H., HANCOCK R.E.W., LOY S., OLSON M.V.;
> RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
> RI opportunistic pathogen.";
> RL Nature 405:959-964(2000);
> DE EMBL: AF004613; AAG052611;
> DR InterPro: IPR001343; HemIysn_Ca_Find.
> DR InterPro: IPR003880; Phosphopant_attach.
> DR EPRINTS: PK0013; CACNCH38T.
> DR PROSITE: PS00012; FHSDFPACTTCHILIL; CNRWR_L1.
> KW Hypothetical protein; Complete proteome.
> SQ SEQUENCE 2458 AA: 23-111 MW: 135564VABDQ18899 CRG64;
>
> Query Match 5.8%, Score 187, DB 3, Length 2468
> Best local similarity: 19.5%, Pval: 7.4e-05;
> Matches: 170, Conservation: 9, Mismatches: 30, Indels: 119, Gaps: 39.
>
> 3 AFISGATGRRMAVTCAGAGNENRATAGCAHMLMHTFANDL .....
> 504 APNAPVVPFNSGSELSTADGSSVT-LITGNGNFIIGTATANANMSTSTSTPEEGIV 562
>
> 52 ..... PHITITMKTKTQNNVT ..... SKIPFGQNNQNNW 83
>
> 56 VVVAVFCAVPSFAYVAVFATVTFMILSTFETSSVTLICNGNR 120 121
>
> 84 RHEVYLSSTGTHM GSVFVAGSWGADSTFSSNEETFAFVY ..... 124
>
> 622 Q VVADGSGNTTTFTRFVNSTVNAATADGSSAGSSASVTVDAVAVTPVNS 677
>
> 125 ..... RVATTFANVFWSTAFVFAVFAVAVSTAFVFCFPMQ PFIDLP 169
>

```

```

> 678 NSTTISCHAHGATVTHHNGNR 120VTA ..... DPGSGNSTTTTITFNR 724
>
> QY 169 ..... IVPAAALEPISGVLMSSYKNDAPGSSPG-GILL 203
>
> DB 725 TVVNATATDASGNTSAGSSVTVDSVAPATVINSNGTTLSTIA ..... EPGSSVTL 776
>
> QY 204 TS ..... SMDPSTGIVSDRTVYTKD ..... 225
>
> DB 777 TDGNGNPVGYVTAAGSGNMSFEPSTPLALGVVNAIATLPAANISAGDSTVNSVAPTP 836
>
> QY 226 --MCPGISMDCNQ---IVTGDNDAKTISLYDSSDSWVLPDGMQVAPGQSSATWS 279
>
> DB 837 TVNISNGSSISGTAEPGSTVITITGNNNPVTAFTADGSGNNTYPTSTPLANGTVVNVAG 846
>
> QY 280 DGRVFTIGSSGSGVFEKNCE---VYFSFKT---MTSLFNKV ..... NFM-LT 322
>
> DB 897 D---AAGNSSPGASVTVDSQAPAPVAVVNSGTTLSGIAEPGATVITLDGNGNPVGYT 952
>
> QY 923 ADKQGLYPSDHNALFPRKRGSVFQAGPSTAMNVTYSr-----SQTV-----KS 367
>
> DB 953 ADGSG-----NMSFTPTPTPLANGTVVNAVATASDPTGNTS 985
>
> QY 968 AGKPSNMGVAHIAACSAVMTAVKSK 113GSGHYQSDR 409
>
> DB 986 APASTIVNSVAPAPVVPNSGASRHSIAEPVATVTLHNSGNPLIGVTALESNNMSTFP 1045
>
> QY 410 ATTAHITTTIGPPTSP-----NTVFASNGVYFAPTFH-TSVVLP 448
>
> DB 1046 STPLADGIVVNAIADPAGNIGQSSIHVDALAPVAVLNSSGSLSGIAEPGSTVITL 1105
>
> QY 449 DGS-----TFTIGSGRPGIPFEDSTPVFTPELY-VPEQTPTKQNNSTVYRVHS---ISL 500
>
> DB 1106 TQNNPLAEVTAISGNNVYVPSPIANGLVNVVAVVADSSNSPFAVIVYDSSAPPADV 1165
>
> QY 501 LLP-DGRVFNQ-----GGGCGPCTTHPDAQITPVLYNSPDLATPKI 546
>
> DB 1166 INPNSVAVISTAFARVATVTLTASGNPLIGVTALESNNMSTFPs-----PLANGTV 1219
>
> QY 547 TRISQVSV-KVQSKVTSISHSISKASIIPVSAIHTVNIHQKIPVITITNNNSYSFQ 605
>
> DB 1220 VAVADFTIGNCHVAAIVLVAATVAVTIDNSCHLISGVACAKVILLDNGNNTGTE 1278
>
> QY 606 VPSDSVALLPQVWMLPVNSAGVSVASII 635
>
> DB 1279 TTADGS---GNWSF---TPCTPLANGTV 1300
>
>
> RESULT 13
> ID Q01773 PRELIMINARY PRT 559 AA.
> AC Q01773,
> DI 01-NOV-1996 (TEMBLEL 01, Created)
> DI 01-NOV-1996 (TEMBLEL 01, Last sequence update)
> DI 01-AUG-1998 (TEMBLEL 07, Last annotation update)
> DE GLYOXAL OXIDASE PRECURSOR.
> GN GLX2.
> OS Phanerochaete chrysosporium.
> OC Eukaryota; Fungi; Basidiomycota; Hymenochytriales;
> OC Aphyllophorales; Corticiaceae; Phanerochaete.
> OX NCBI_taxid=5906,
> RN 111
>
> SEQUENCE FROM N A.
> KA Kersten P., Witte C., Vanden Wymelenberg A., Cullen D.;
> FL J. Bacteriol. 0:0-0(0).
> DE EMBL: L47287; AAA87595.1;
> KW signal.
> FT SIGNAL 1 22
> FT CHAIN 23 559 GLYOXAL OXIDASE.
> SQ SEQUENCE 559 AA: 59140 MW: 5145PAI14298866 CRG64;
>
> Query Match 5.3%, Score 182.5, DB 3, Length 559;
> Best local similarity: 21.8%, Pval: 0.0001;
>

```

139 STAEINVEASSTAP-----QPGIGFWPTIDPIVPAATAEPTSGVLMSSVRNDAF 194
139 STAEINVEASSTAP-----QPGIGFWPTIDPIVPAATAEPTSGVLMSSVRNDAF 194
3 SLAAVSLAATAAIAAASLAIQWELIKINLSIYALIAIVNSSLIVIF---EFAT 57
196 GGNP-----GGTILSSNIPISGIVSDRTVITKHEMFCGISMONGQIVVIGNDA--- 247
58 GDPLKINGESTWGLMDLDTIVRLSVLT---DSFASGALLNGITWVSGGTGGTG 114
248 -----KRTSLYSSSDSWIPGPD-----MCVAPGYSSATMSDGPVFTI 286
115 GVAAPPGNCAITPFPQAS---PSGIGCTILFEDPAIVHLIFKWWYSSVRIPDSLMII 171
287 GGSWSGVFPKNGEVSF-----SKTWTSLPNAKVPMLTAKGCIYRSND 333
172 GGS-----HVLIPFYVNDPANSFEFPFSKEQTPRPSAFLESLAN---LEPR--- 216
234 HAMLFWKKGVSFGAGSTAMNWTYSGSGVKSAGKPGSNPVAPD-----AMCG 384
217 ---AFALPDGVFVANNOSITY-----DIE---KNETILIPDPNGVPTNPIDG 261
185 NAVW DAVKGLITFGSGSDYQSDATNAHITLGEPTG-----PNTVA 431
262 SAILPLSPDPPI-EVLVCGSS---TAPISPTSLSSQHPATSSPRTITPEGIKA 416
432 S---NGTPARTPITSVLDGSGFTTGGGPPGI-----PREDST---PVPEPEIYV 477
417 GWVHEHLEARMDELHVPGQILITNGAGTGFALSAVADPGNSNADHVLPLSLYT 376
478 FEQDTYKQNTN---SIVVYHSISLLDFGVNMGQ-----GCGGCTTNNH 522
477 FDA-PLGKRTSNACMPITLIPRHSIVITLQGNPFQGNPNMNFPPGIGIKPFSF 435
523 FDAQITPNTLYNSGNLAPPKTPTTSTQSVKVG---TITSTDSISK-ASLIRY 576
436 LRLEIDLPPEPKSPALTIMP-----EKLKFGQKVTVDIIPDLKASKVOVALMDL 488
577 GVAHTVNTDQRIPLTLTNNNG-NSYSFGVPSDGVALPGVMMLF-----VM 623
489 GFSHAFHSSAFVFMSSISADPKSITFTAPPNCPVTPPPAVVFTITIVTSRGEVW 548
624 NSAGVP 629
549 MGSQNP 554

PM: 118991, AAA3747.L, -
KW Signal: Peroxidase, 22
FT SIGNAL 23 559 POTENTIAL
SC SEQUENCE 559 AA, 5916 MW, 944-PVAL-101906095 GRC64;
Query Match 53%; Score 192.5; P: 3; Length 559;
Best local similarity 21.8%, Prod 37.00001;
Matches 132, Conserved 87, Mismatches 229, Indels 169, Gaps 30;
QY 139 STAEINVEASSTAP-----QPGIGFWPTIDPIVPAATAEPTSGVLMSSVRNDAF 194
Db 3 SLAAVSLAATAAIAAASLAIQWELIKINLSIYALIAIVNSSLIVIF---EFAT 57
QY 196 GGNP-----GGTILSSNIPISGIVSDRTVITKHEMFCGISMONGQIVVIGNDA--- 247
Db 58 GDPLKINGESTWGLMDLDTIVRLSVLT---DSFASGALLNGITWVSGGTGGTG 114
QY 248 -----KRTSLYSSSDSWIPGPD-----MCVAPGYSSATMSDGPVFTI 286
Db 115 GVAAPPGNCAITPFPQAS---PSGIGCTILFEDPAIVHLIFKWWYSSVRIPDSLMII 171
QY 287 GGSWSGVFPKNGEVSF-----SKTWTSLPNAKVPMLTAKGCIYRSND 333
Db 172 GGS-----HVLIPFYVNDPANSFEFPFSKEQTPRPSAFLESLAN---LEPR--- 216
QY 234 HAMLFWKKGVSFGAGSTAMNWTYSGSGVKSAGKPGSNPVAPD-----AMCG 384
Db 217 ---AFALPDGVFVANNOSITY-----DIE---KNETILIPDPNGVPTNPIDG 261
QY 185 NAVW DAVKGLITFGSGSDYQSDATNAHITLGEPTG-----PNTVA 431
Db 262 SAILPLSPDPPI-EVLVCGSS---TAPISPTSLSSQHPATSSPRTITPEGIKA 416
QY 432 S---NGTPARTPITSVLDGSGFTTGGGPPGI-----PREDST---PVPEPEIYV 477
Db 417 GWVHEHLEARMDELHVPGQILITNGAGTGFALSAVADPGNSNADHVLPLSLYT 376
QY 478 FEQDTYKQNTN---SIVVYHSISLLDFGVNMGQ-----GCGGCTTNNH 522
Db 477 FDA-PLGKRTSNACMPITLIPRHSIVITLQGNPFQGNPNMNFPPGIGIKPFSF 435
QY 523 FDAQITPNTLYNSGNLAPPKTPTTSTQSVKVG---TITSTDSISK-ASLIRY 576
Db 436 LRLEIDLPPEPKSPALTIMP-----EKLKFGQKVTVDIIPDLKASKVOVALMDL 488
QY 577 GVAHTVNTDQRIPLTLTNNNG-NSYSFGVPSDGVALPGVMMLF-----VM 623
Db 489 GFSHAFHSSAFVFMSSISADPKSITFTAPPNCPVTPPPAVVFTITIVTSRGEVW 548
QY 624 NSAGVP 629
Db 549 MGSQNP 554

RESULT 15
Q9GCEU1 PRELIMINARY; PRT; 1441 AA.
ID Q9GCEU1
AC Q9GCEU1
DT 01-JUN-2001 (TEMBREL 17, Created)
DT 01-JUN-2001 (TEMBREL 17, Last sequence update)
DT 01-JUN-2001 (TEMBREL 17, Last annotation update)
DE UNKNOWN PROTEIN
GN VOIC
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis)
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OX Lactococcus
ON NCBI_TaxID=1360;
RX 111
PP SEQUENCE FROM N.A.
PC STRAIN-T11403;
RA Pothier A., Wincker P., Manger S., Tallon O., Matarne K.,

GenCore version 4.5
Copyright 1993-2000 GenCore Inc.

us-09-722-602-10

Protein search using SW Model
Protein: 1016,226 Million cell updates/sec
(without alignments)

Parent: 1016,226 Million cell updates/sec
Sequence: 1016,226 Million cell updates/sec

Scoring table:
Gap: 10.0 - Gap: 0.5

Search: 21252 acgs, 2259292 residues

Total number of hits satisfying chosen parameters: 21252

Post processing: Minimum Match 100%
Listing first 45 summaries

Database: 1 Issued Patents: AA*
2. 1016,226 Million cell updates/sec
3. 1016,226 Million cell updates/sec
4. 1016,226 Million cell updates/sec
5. 1016,226 Million cell updates/sec
6. 1016,226 Million cell updates/sec

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2207	64.4	US-09-722-602-10	Sequence 2, Appl
2	2207	64.4	US-09-722-602-10	Sequence 2, Appl
3	131	129	US-09-722-602-10	Sequence 2, Appl
4	130.5	3.8	US-09-722-602-10	Sequence 2, Appl
5	130	3.8	US-09-722-602-10	Sequence 2, Appl
6	129	3.8	US-09-722-602-10	Sequence 2, Appl
7	129	3.8	US-09-722-602-10	Sequence 2, Appl
8	127.5	3.7	US-09-722-602-10	Sequence 2, Appl
9	127.5	3.7	US-09-722-602-10	Sequence 2, Appl
10	125.5	3.7	US-09-722-602-10	Sequence 2, Appl
11	125.5	3.7	US-09-722-602-10	Sequence 2, Appl
12	125.5	3.7	US-09-722-602-10	Sequence 2, Appl
13	125.5	3.7	US-09-722-602-10	Sequence 2, Appl
14	125.5	3.7	US-09-722-602-10	Sequence 2, Appl
15	125	3.6	US-09-722-602-10	Sequence 2, Appl
16	125	3.6	US-09-722-602-10	Sequence 2, Appl
17	123.5	3.6	US-09-722-602-10	Sequence 2, Appl
18	123.5	3.6	US-09-722-602-10	Sequence 2, Appl
19	123	3.6	US-09-722-602-10	Sequence 2, Appl
20	123	3.6	US-09-722-602-10	Sequence 2, Appl
21	123	3.6	US-09-722-602-10	Sequence 2, Appl
22	123	3.6	US-09-722-602-10	Sequence 2, Appl
23	123	3.6	US-09-722-602-10	Sequence 2, Appl
24	123	3.6	US-09-722-602-10	Sequence 2, Appl
25	123	3.6	US-09-722-602-10	Sequence 2, Appl
26	123	3.6	US-09-722-602-10	Sequence 2, Appl
27	123	3.6	US-09-722-602-10	Sequence 2, Appl

28	122	3.5	US-08-722-602-10	Sequence 2, Appl
29	122	3.5	US-08-722-602-10	Sequence 2, Appl
30	120	3.5	US-08-722-602-10	Sequence 2, Appl
31	119	3.5	US-08-722-602-10	Sequence 2, Appl
32	119	3.5	US-08-722-602-10	Sequence 2, Appl
33	119	3.5	US-08-722-602-10	Sequence 2, Appl
34	119	3.5	US-08-722-602-10	Sequence 2, Appl
35	118	3.4	US-08-722-602-10	Sequence 2, Appl
36	118	3.4	US-08-722-602-10	Sequence 2, Appl
37	117	3.4	US-08-722-602-10	Sequence 2, Appl
38	116.5	3.4	US-08-722-602-10	Sequence 2, Appl
39	116.5	3.4	US-08-722-602-10	Sequence 2, Appl
40	115	3.4	US-08-722-602-10	Sequence 2, Appl
41	115	3.4	US-08-722-602-10	Sequence 2, Appl
42	115	3.4	US-08-722-602-10	Sequence 2, Appl
43	114.5	3.3	US-08-722-602-10	Sequence 2, Appl
44	114.5	3.3	US-08-722-602-10	Sequence 2, Appl
45	113.5	3.3	US-08-722-602-10	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-257-536-2
Sequence 2, Application US-09-257-536
Patent No. 6090604
GENERAL INFORMATION:
APPLICANT: Golightly, Elizabeth
APPLICANT: Berka, Randy M.
APPLICANT: Foy, Michael W.
TITLE OF INVENTION: Polypeptides Having Salts of Sulfate
FILE REFERENCE: 582, 000 US
PARENT APPLICATION NUMBER: 58/000,000
CURRENT FILING DATE: 1999-02-24
NUMBER OF SEQ. ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ. ID NO. 2
LENGTH: 679
ORGANISM: Fusarium
TYPE: PRT
US-09-257-536-2

Query Match	64.4%	Score 2207	DB 3	Length 679
Best Local Similarity	64.0%	Prod. No. 26-178		
Matches 407	Conservative 83	Mismatches 146	Indels 10	Gaps 8
1	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
2	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
3	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
4	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
5	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
6	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
7	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
8	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
9	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
10	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
11	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
12	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
13	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
14	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
15	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
16	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
17	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
18	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
19	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
20	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
21	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
22	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
23	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
24	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
25	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
26	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
27	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
28	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
29	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
30	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
31	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
32	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
33	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
34	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
35	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
36	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
37	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
38	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
39	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
40	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
41	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
42	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
43	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
44	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		
45	ASAPLIG-ALSPNNMAVIGUSAGSSEENKALHKKKJFHHFYSAGNDPPPHIYILUM	59		

PARALLEL FILING DATE: 1998-07-28
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ ID NO 2

LENGTH: 915
 TYPE: PRT
 ORGANISM: Bacillus deramificans
 FEATURE:
 NAME/KEY: PEPTIDE
 LOCATION: (1) - (915)
 OTHER INFORMATION: Pullulanase
 S: 7-349-257 2

Query Match 3.8%; Score 130, DB 4, length 915;
 Best Local Similarity 19.4%; Pred No. 0.021;

Matches 129, Conservative 63, Mismatches 189, Indels 254, Gaps 33;

192 DAFG...GSE...GILLISSMSTGLVSOHLVIVK... 223
 42 LSEGVASADIPENPSQVGLVRIQEM IKDVSADRYIDLKSGNEVLEQNSQIFENE 99
 424 ---HMFPGIS---MDSGQIVV-----TGNDAKKTLTSS- 256
 130 KPAELAAVAVASAVIASMOVAVNISQPLTSGASGTVHATANKDLFVTVKASL 159
 257 ---SDSWIPDQVAR-----GYSSATMS--- 279
 160 GGVAVAVANTTCHTGGKATDINTELEKVTNRICVRSCHLFDGVYKVMASNO 219
 280 ---GKAVTTGSSSSG GYTKRGLVSTPS TWISLPAKAVPMKID 324
 420 PAVSVAVMGRFPLGATWLMQAV NEALYSSGVKTDLVVTLGEGFDVSELTSG 276
 425 KQJ-----LYKSDN-----HAW-----LEPMKQGV 345
 277 IIVYCAKAVYKRVNLSSQYTSIRAGHTYQKATIKKVAWLTSTQVNVLLTDSATQSV 336
 446 PVAIDPSTAM-----NMVY---TSGSDYKSAKQSNQVAFDAMGNAM 388
 432 TKLVHMASHEVMAIVNLENWYMYEVLTGGSTRTA-----VDPA--TAIA 385
 439 YAAVEFELIFSSSEYLSIA---TNAH-----ITLCPGSGSPRVIVASN 433
 486 ENJIRKIV---DLAKTDPAWNSDKHTPENIDEVYEMVIVRPSIDPNSGMKK 439
 434 GUYFARFITSVAVLPDS-STFTTGGGRGIDPEEDSTVFT-----PEI 475
 440 GKYIATL-EKGIEKPPNVKRTIOSI KQIGITHVGLMVPVANSVDPTDPQDNMGYDPRN 498
 476 Y-VEGDEITYKPNPNSIVRVYHSITLLPDSVEFNSSGSLGKCTTH-FAAQI-----F 528
 499 YVPEHGVYVAINAMNAEIKKEKREVLSTLREHI-----JVMQDVVYVNHFTFQISDFPKI 553
 529 IINYVYNSD-----GN-LATRPKLITRSTOSVKV-----GGRTITSTUSS 567
 554 VEEYTRIRIDAGNYTNGSGTNEIAAEKRPVQKTI DLSLKYWVNEHIDGRFDLM--AL 611
 568 ISKASTIRYCATHTVN 584
 612 ICKKMSKASHHAIN 628

RES: 1-157A-5
 Sequence 5, Application US/0731157A
 Patent No. 5457032

GENERAL INFORMATION:
 APPLICANT: Onaka, Wilhelmus J.
 APPLICANT: Missel, Onno
 APPLICANT: Van der Laan, Jan M.
 APPLICANT: Lottius, Herman B.M.
 TITLE OF INVENTION: Mutated beta-lactam acylase genes

NUMBER OF SEQUENCES: 50
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: COOLEY GODWARD CASTRO HODDLESON & TATUM
 STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
 CITY: PALO ALTO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94306

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 09/0731157A
 FILING DATE: 1990509
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PP 60200962
 FILING DATE: 18-APR-1990

ATTORNEY/AGENT INFORMATION:
 NAME: RAE-VENTER P.H.D., BARBARA
 REGISTRATION NUMBER: 32,750
 REFERENCE/SECRET NUMBER: 7890 027 0005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-494-7622
 TELEFAX: 415-857-0663

TELEX: 380816 COOLEY PA
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 846 amino acids
 TYPE: AMINO ACID
 TOPLOGY: linear

MOLECULE TYPE: protein
 ORIGINAT SOURCE:
 ORGANISM: Escherichia coli
 PURIFICATION INFORMATION:
 AUTHORS: Schumacher, G

AUTHORS: Sizmann, D
 AUTHORS: Haug, H
 AUTHORS: Buckel, P
 AUTHORS: Bock, A
 TITLE: Penicillin acylase from E.coli: unique
 TITLE: gene-protein relation.
 JOURNAL: Nucleic Acids Res.
 VOLUME: 14
 PAGES: 5713-5727
 DATE: 1986

US-07-731-157A-5

Query Match 3.8%; Score 129, DB 1, length 846;
 Best Local Similarity 19.0%; Pred. No. 0.022;

Matches 132, Conservative 63, Mismatches 241, Indels 258, Gaps 40;

2 SATGSAISRNMAVATPDSAGSNEC-----NKALDQ-----NKDT 37
 211 SAPTTAVQESNPLPKFNQNSQITALLPRYDLPAPMLDEPAGADGALLATGAKRRET 270
 38 FWHTEY--GAN---GDKPKPHYTIDMKTTQVNVNGLSMLPKQDQNGW-----IGK 84
 271 IVOGFQCCANGLAGYPTISNMVIGSKRAQIAKAIWV NCPQGWAVAPATYVIGL 326
 85 HEVYLSDDG-----NMGSPVA-----SSSWF 106
 327 HCAGYDVNTGTPAYGVVGHNGVLSWSTIAGFGVVDIFARKISAEKQVYIIMQKVV 386
 107 ---ADSTIKYSNFEIRP---ARYVRLVALL- 131
 367 KMLSRREITLVKNGQAEFTTVMTVGHNLQITQDTQTAVAKSRAWGKEVASILAMTHQ 446
 132 ---ANQPMTSAEINVPQASSYA-----PQRTGAWGP 163

Db 447 MKAPRMVWVQAAVGLTITRYVACGNNTIVITCTITPQSHRRLIVCTKRWK 506
 QY 164 TIDPLVAAAALPTSGVLMSSY-----RNDAPGSPMLITLSSWDPTGIVSR 217
 Db 507 KATPTEKREKVRINQSTIAWNNKLNENLAALAWNA+EQ
 QY 219 IVYTRKPKPPTKPPKPNLVTVGKAKETSYSSSTSTPTKPTKVAEYVSAT 277
 Db 553 VTEIDPTEKPPETLAWAVI-----KQTSQINIKH-----SGL 600
 QY 279 MSDEK VETLQANQSVIENHVEVLEERTVETLAAVHGLAAQLAVEDNE 334
 Db 601 WSDPKQVLEL-TRWDG-----ILNDEKRIKQKPSAII 637
 QY 435 AMLETKRSVIGQSTANWVY 532SVASAPSSNRYVALANQAAVAVK 354
 Db 638 WLTSMLEKRIYAAVPRPQWYASQYIISQPSINISV-----DAVILYAVQ 691
 QY 395 -----KILFGSSPDVQSDATINAHITLCEPRTSINTVYASGLTARTHTS--V 445
 Db 692 TESTPFAVEEACHTGIVLAALEQWLEEE PKNVSNMPTMALTPFAHP 749
 QY 446 VLPDSTFEEG-----GEGG-----PEPESL IVEHLAVLGLIYKQKSLV 493
 Db 750 GVPQAAELTHQAEYQKQTEENEMIVPTTSRPVLAHCVAVPQSGF 799
 QY 194 VYHSTSLITPQGVFNQSGDQCTTAPPAQI 527
 Db 800 -----IAPDQIV-----DKHYEDQL 814

RESULT 7

OS-08-541-780-5
 Sequence 5, Application OS/06541780
 Patent No. 5935831

GENERAL INFORMATION:
 APPLICANT: Quak, Wilhelmus J.
 APPLICANT: Misset, Onno
 APPLICANT: Van der Laan, Jan M.
 APPLICANT: Leendert Herman B M.
 TITLE OF INVENTION: Mutated beta-lactam acylase genes
 NUMBER OF SEQUENCES: 50
 CORRESPONDENCE ADDRESS:
 ADDRESS: COFFEY GILBERT CASHE/RODFERSON & TATUM
 STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
 CITY: PALO ALTO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94306
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: WINDOWS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25 (EP0)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/06541 780
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 06/07731177
 FILING DATE:
 APPLICATION NUMBER: EP 0020062
 FILING DATE: 18-APR-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: RAY VENTER PH.D., BARBARA
 REGISTRATION NUMBER: 42 750
 REFERENCE: 427 00003 396 027 00003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-494-7422
 TELEFAX: 415-857-6663
 TELEX: 340816 COFFEY PA
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:

LENGTH: 846 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Escherichia coli
 PHYLATION INFORMATION:
 AUTHORS: Schumacher, G
 AUTHORS: Szymacher, G
 AUTHORS: Haug, H
 AUTHORS: Haug, H
 AUTHORS: Buckel, P
 TITLE: Penicillin acylase from E. coli: unique
 TITLE: gene-protein relation.
 JOURNAL: Nucleic Acids Res.
 VOLUME: 14
 PAGES: 5713 5727
 DATE: 1986
 OS-08-541-780-5

Query Match 3.8%, Score 129, DB 2, Length 846,

Best Local Similarity 19.0%, Ident. No. 0.023,
Matches 132, Conservative 63, Mismatches 241, Indels 256, Gaps 30;

QY 2 SAPLGSAISKNNMAVLDSSAOSGNEC-----NKALDG-----NKDT 37
 Db 211 SAPTIAVQESNYPLKFNQNSQTAALPPYTPAPMTTPPAKGMATIALATARKNPET 270
 QY 38 FWHTEY--GAN--GDPRKPHYTITDKTKTONVNGLSMLPRQAGNWNW-----IGR 84
 Db 271 IVAQFAUGGANGIAAYPTTSMNMTGKSKAQDAKAIW-----NQTQFGYALAYTYGDL 326
 QY 85 HEVYLSISQI-----NMQSPVA-----SGSW 106
 Db 327 HGAGYEVGNTPFAVPPLVFGHGVISWSTAGFQDPVDIFAEPLSAEKPGYLLHNGKV 386
 QY 107 -----ADSTKYSNFEETP--ARYRVIAITE- 131
 Db 347 KMLSPPELITVKNQJAEETVWQVYVH;NLTQVQITVIAVASKAWQKQEVASLAWHQ 446
 QY 132 ---ANGPMTSLAEINVAQSSYVA-----PQPTGKMP 163
 Db 447 MKANMWLEWLEAAKQDALIMVYALVW;NIGVHT;AAYPLQSSM;PRLVPGTGMKW 506
 QY 164 TIDPLVAAAALPTSGVLMSSY-----KNDAPGSPMLITLSSWDPTGIVSR 217
 Db 507 KGLPFEKNRPVYNPQSGYIANMNNSPQKDYPAIDLFAIWMG-----ADR 552
 QY 218 IVYTRKPKPPTKPPKPNLVTVGKAKETSLYDSSSDSMIFQPDQVARGYSSAT 277
 Db 553 VTEIDPTEKPPETLAWAVI-----KQTSQINIKH-----SGL 600
 QY 279 MSDEK VETLQANQSVIENHVEVLEERTVETLAAVHGLAAQLAVEDNE 334
 Db 601 WSDPKQVLEL-TRWDG-----ILNDEKRIKQKPSAII 637
 QY 435 AMLETKRSVIGQSTANWVY 532SVASAPSSNRYVALANQAAVAVK 354
 Db 638 WLTSMLEKRIYAAVPRPQWYASQYIISQPSINISV-----DAVILYAVQ 691
 QY 395 -----KILFGSSPDVQSDATINAHITLCEPRTSINTVYASGLTARTHTS--V 445
 Db 692 TESTPFAVEEACHTGIVLAALEQWLEEE PKNVSNMPTMALTPFAHP 749
 QY 446 VLPDSTFEEG-----GEGG-----PEPESL IVEHLAVLGLIYKQKSLV 493
 Db 750 GVPQAAELTHQAEYQKQTEENEMIVPTTSRPVLAHCVAVPQSGF----- 799
 QY 194 VYHSTSLITPQGVFNQSGDQCTTAPPAQI 527
 Db 800 -----IAPDQIV-----DKHYEDQL 814

Sequence 2: Application US/09313677
Patent No. 6400115
GENERAL INFORMATION:
APPLICANT: Teague, W. Martin
APPLICANT: Brumm, Phillip J.
APPLICANT: Allen, Larry N.
APPLICANT: Brinkman, Igor A.
TITLE OF INVENTION: Amylase Promoter and Leader Sequences
FILE REFERENCE: Pullulanase Constructs
CURRENT APPLICATION NUMBER: US/09/313,677
EARLIER FILING DATE: 1999-05-18
NUMBER OF SEQ ID NOS: 22
SEQUENCE: Patent In Ver. 2.0

Query Match 3.7%; Score 127.5; DB 4; Length 946;
Best Local Similarity 20.1%; Pred. No. 0.034;
Matches 147; Conservative 68; Mismatches 241; Indels 277; Gaps 39;
75 DGNNGMTGHEHYLSSDGTNM-----GSP--- 99
1 DGNNTN-IVVHYFRPSDGYDWMNMMWPFENDGALYDFNPTDSYGEVASYDIPGNPSOV 59
100 ---VASNSMFADSTTKYSNFTPRARYRLV-----ATTPANQPMSTIAEINVF 146
40 GIVVKNQWIAADISKIDLSKSHETWVQNSQITSTKJAEAAQPAVSNAYLNAS 119
147 QASSYIAPQP---GLGHWGPTI-----DLPIVFAAAALEPT----- 179
120 NLYVYKISQPTLQPSGSGFTVHDTANKDIPVTSVDANQVAVIAGTFHFGGSMA 179
160 ---SC-----KVLMSSTYKNAFAFGS-----PGGTTITS 205
169 PGNHMLLEKVENELYEFSDNPEPNYVFAVAKNSWNPSPSPNNTNIVPAGAHYTF 239
206 SWDSTGIVSDRIVTYVKHMPFGGISMDNG---QIVVTGMDAKKTSLYDSSDSWI 261
240 SYLPTSHAVVD-TINNNDAD-----LQVDSGCVKTDIAVAVTLGENP-----DVSHTLSI 287
262 PGDPMVAPPYGSSATMSDGFVFTTIGSSWSGCVKKNQVYSSTSNLISLPMNAVIM 321
288 QTEHYQA---GVVLPKRVLSQVYYSQDIAQNTVTKNATTF-----KVWAP-TSTOVNVL- 339
442 TARKQGYKSRNHMLPGWKGSGVQAGPSTAM-----NMY-----TSGSD 364
440 ---LYNS-----ATGAVYKIVPMTASGHWAEATVNDLENNYMYEVTGGSS 384
465 VESAGPESQSNFVAPFMAGNAVMYDAVFKITTPGSS-PLYQDSDATTAH 416
485 TRTA-----VQFYA---TALAHNTHGMIWIAKTDPAQWESD-----KHITPKNIED 429
417 ---TIGPGTSPNIVFASNGIYFARFHTSVVLPLPG-STFTIGGQPGPIPERSTIV 470
480 EYVIMVVEFVFNSSGMRKPGYVATL-EPGTFVPMVKTAVSLKQGITTHVQLQPV 488
471 FT-----PELY-VPEGTFKQNPNSIVH--YHSISLLPQDGVFNG 510
484 FAKNSVNLNLTQYRWQYDPRNNTFNG--YATNANCTTKKEKEEMVSIHQDH 543
511 GAGTCTCTNN-FDAQT-----FTPNVLYNSD-----GN--LATREKTRIST 551
544 ---VNMIVVYNNHTATQISDEKIVIEVYVPTDAGNTNGSGTGNELAAEPKPVKFTI 601

Query Match 3.7%; Score 127.5; DB 4; Length 946;
Best Local Similarity 20.1%; Pred. No. 0.034;
Matches 147; Conservative 68; Mismatches 241; Indels 277; Gaps 39;
75 DGNNGMTGHEHYLSSDGTNM-----GSP--- 99
1 DGNNTN-IVVHYFRPSDGYDWMNMMWPFENDGALYDFNPTDSYGEVASYDIPGNPSOV 59
100 ---VASNSMFADSTTKYSNFTPRARYRLV-----ATTPANQPMSTIAEINVF 146
40 GIVVKNQWIAADISKIDLSKSHETWVQNSQITSTKJAEAAQPAVSNAYLNAS 119
147 QASSYIAPQP---GLGHWGPTI-----DLPIVFAAAALEPT----- 179
120 NLYVYKISQPTLQPSGSGFTVHDTANKDIPVTSVDANQVAVIAGTFHFGGSMA 179
160 ---SC-----KVLMSSTYKNAFAFGS-----PGGTTITS 205
169 PGNHMLLEKVENELYEFSDNPEPNYVFAVAKNSWNPSPSPNNTNIVPAGAHYTF 239
206 SWDSTGIVSDRIVTYVKHMPFGGISMDNG---QIVVTGMDAKKTSLYDSSDSWI 261
240 SYLPTSHAVVD-TINNNDAD-----LQVDSGCVKTDIAVAVTLGENP-----DVSHTLSI 287
262 PGDPMVAPPYGSSATMSDGFVFTTIGSSWSGCVKKNQVYSSTSNLISLPMNAVIM 321
288 QTEHYQA---GVVLPKRVLSQVYYSQDIAQNTVTKNATTF-----KVWAP-TSTOVNVL- 339
442 TARKQGYKSRNHMLPGWKGSGVQAGPSTAM-----NMY-----TSGSD 364
440 ---LYNS-----ATGAVYKIVPMTASGHWAEATVNDLENNYMYEVTGGSS 384
465 VESAGPESQSNFVAPFMAGNAVMYDAVFKITTPGSS-PLYQDSDATTAH 416
485 TRTA-----VQFYA---TALAHNTHGMIWIAKTDPAQWESD-----KHITPKNIED 429
417 ---TIGPGTSPNIVFASNGIYFARFHTSVVLPLPG-STFTIGGQPGPIPERSTIV 470
480 EYVIMVVEFVFNSSGMRKPGYVATL-EPGTFVPMVKTAVSLKQGITTHVQLQPV 488
471 FT-----PELY-VPEGTFKQNPNSIVH--YHSISLLPQDGVFNG 510
484 FAKNSVNLNLTQYRWQYDPRNNTFNG--YATNANCTTKKEKEEMVSIHQDH 543
511 GAGTCTCTNN-FDAQT-----FTPNVLYNSD-----GN--LATREKTRIST 551
544 ---VNMIVVYNNHTATQISDEKIVIEVYVPTDAGNTNGSGTGNELAAEPKPVKFTI 601

Sequence 19: Application US/09313677
Patent No. 6300115
GENERAL INFORMATION:
APPLICANT: Teague, W. Martin
APPLICANT: Brumm, Phillip J.
APPLICANT: Allen, Larry N.
APPLICANT: Brinkman, Igor A.
TITLE OF INVENTION: Amylase Promoter and Leader Sequences
FILE REFERENCE: Pullulanase Constructs
CURRENT APPLICATION NUMBER: US/09/313,677
EARLIER FILING DATE: 1999-05-18
NUMBER OF SEQ ID NOS: 22
SEQUENCE: Patent In Ver. 2.0

Query Match 3.7%; Score 127.5; DB 4; Length 943;
Best Local Similarity 20.1%; Pred. No. 0.034;
Matches 147; Conservative 68; Mismatches 241; Indels 277; Gaps 39;
75 DGNNGMTGHEHYLSSDGTNM-----GSP--- 99
1 DGNNTN-IVVHYFRPSDGYDWMNMMWPFENDGALYDFNPTDSYGEVASYDIPGNPSOV 66
100 ---VASNSMFADSTTKYSNFTPRARYRLV-----ATTPANQPMSTIAEINVF 146
67 GIVVKNQWIAADISKIDLSKSHETWVQNSQITSTKJAEAAQPAVSNAYLNAS 126
147 QASSYIAPQP---GLGHWGPTI-----DLPIVFAAAALEPT----- 179
127 NOVIVKISQPTLQPSGSGFTVHDTANKDIPVTSVDANQVAVIAGTFHFGGSMA 186
180 ---SC-----KVLMSSTYKNAFAFGS-----PGGTTITS 205
187 PGNHMLLEKVENELYEFSDNPEPNYVFAVAKNSWNPSPSPNNTNIVPAGAHYTF 246
206 SWDSTGIVSDRIVTYVKHMPFGGISMDNG---QIVVTGMDAKKTSLYDSSDSWI 261
247 SYLPTSHAVVD-TINNNDAD-----LQVDSGCVKTDIAVAVTLGENP-----DVSHTLSI 294
262 PGDPMVAPPYGSSATMSDGFVFTTIGSSWSGCVKKNQVYSSTSNLISLPMNAVIM 321
295 QTEHYQA---GVVLPKRVLSQVYYSQDIAQNTVTKNATTF-----KVWAP-TSTOVNVL- 346
442 TARKQGYKSRNHMLPGWKGSGVQAGPSTAM-----NMY-----TSGSD 364
440 ---LYNS-----ATGAVYKIVPMTASGHWAEATVNDLENNYMYEVTGGSS 391
471 FT-----PELY-VPEGTFKQNPNSIVH--YHSISLLPQDGVFNG 416
484 FAKNSVNLNLTQYRWQYDPRNNTFNG--YATNANCTTKKEKEEMVSIHQDH 488
511 GAGTCTCTNN-FDAQT-----FTPNVLYNSD-----GN--LATREKTRIST 551
544 ---VNMIVVYNNHTATQISDEKIVIEVYVPTDAGNTNGSGTGNELAAEPKPVKFTI 601

Query Match 3.7%; Score 127.5; DB 4; Length 943;
Best Local Similarity 20.1%; Pred. No. 0.034;
Matches 147; Conservative 68; Mismatches 241; Indels 277; Gaps 39;
75 DGNNGMTGHEHYLSSDGTNM-----GSP--- 99
1 DGNNTN-IVVHYFRPSDGYDWMNMMWPFENDGALYDFNPTDSYGEVASYDIPGNPSOV 66
100 ---VASNSMFADSTTKYSNFTPRARYRLV-----ATTPANQPMSTIAEINVF 146
67 GIVVKNQWIAADISKIDLSKSHETWVQNSQITSTKJAEAAQPAVSNAYLNAS 126
147 QASSYIAPQP---GLGHWGPTI-----DLPIVFAAAALEPT----- 179
127 NOVIVKISQPTLQPSGSGFTVHDTANKDIPVTSVDANQVAVIAGTFHFGGSMA 186
180 ---SC-----KVLMSSTYKNAFAFGS-----PGGTTITS 205
187 PGNHMLLEKVENELYEFSDNPEPNYVFAVAKNSWNPSPSPNNTNIVPAGAHYTF 246
206 SWDSTGIVSDRIVTYVKHMPFGGISMDNG---QIVVTGMDAKKTSLYDSSDSWI 261
247 SYLPTSHAVVD-TINNNDAD-----LQVDSGCVKTDIAVAVTLGENP-----DVSHTLSI 294
262 PGDPMVAPPYGSSATMSDGFVFTTIGSSWSGCVKKNQVYSSTSNLISLPMNAVIM 321
295 QTEHYQA---GVVLPKRVLSQVYYSQDIAQNTVTKNATTF-----KVWAP-TSTOVNVL- 346
442 TARKQGYKSRNHMLPGWKGSGVQAGPSTAM-----NMY-----TSGSD 364
440 ---LYNS-----ATGAVYKIVPMTASGHWAEATVNDLENNYMYEVTGGSS 391
471 FT-----PELY-VPEGTFKQNPNSIVH--YHSISLLPQDGVFNG 416
484 FAKNSVNLNLTQYRWQYDPRNNTFNG--YATNANCTTKKEKEEMVSIHQDH 488
511 GAGTCTCTNN-FDAQT-----FTPNVLYNSD-----GN--LATREKTRIST 551
544 ---VNMIVVYNNHTATQISDEKIVIEVYVPTDAGNTNGSGTGNELAAEPKPVKFTI 601

```

; LENGTH: 1732 amino acids
; TYPE: amino acid

```


[illegible]

```

Db 1396 LTVATK GAVVAGGPNVATDTAAAG-----TNANR-----1339
G2 1416 QVASSV AFGVGLSHWGI LILFLVIA AAALEIFSGVLMSSYFPAAGQSGGIIIS 205
Db 1340 -----OKVVVKS-----NQAITE 1354
G2 236 SMQSTIVSGFVYTRHMFEGISMGQNTIV-----TQNVAKTS-----251
L1 1335 LAKKVAIVGVAAIMAAIF-----VAVGECALILFETLEGAALAKALILLAG 1411
G2 252 -----LYNSSTSTHFEVAVAPGYSATMGTCGVPTIGGSMGCVVFNNGEYVSS 306
L1 1412 KLVYKKGQGNITFALANDSV-----KSAIVGKLLSLQNKKNVNDI SGRKGNFARD 1466
G2 307 KTW-----ISLPNAKVNPMIAKQGLYRSTNAHLEFGWRKDS-----VYLAG- 349
Db 1467 KITGDANIHLNGIASLITDILLNSGATINLGNGITDN-----KKAAASVKQVLNMG 1520
G2 350 -----PSTANMY-----YISGSQVKSAGKPGQSNQVAPDAMG--NAVW 388
L1 1521 NVEGVYASANNVEVLEVAIVGVVHVSDEKRI TSIVTEKINQKPEVYFAKTSVI 1580
G2 389 YDAKKGLITGSGPEVQDSDATNNHIIIEEGPGSPNIVASNLTAFTPHSV--- 445
L1 1581 KIH NGRFET---GRI KIANNN---GVYTHIYKRE-NRIVAKAVITAVNA 1628
G2 446 -VLEGS-----FTGSGEGIPHESTPEFTPIVFEFTVYKQNSIVRVVHS 497
Db 1629 GMPVKYTSANNQNDPFAVSGTNVTRFAGNGIT-TAFV-----TANQSGITVKT- 1677
G2 498 ISLLPGGPFNGSGGGLGPGCTNHHPDQITPN-----YLYNSQNI ATPKTRITST- 551
L1 1678 -NVVVAQKILDEKRIADTIVLVAGQVIAVNNQXKREKMSIASLADENKLSXTATA 1736
G2 552 -----GSKVSGPLTISTISSIKASLIPYGIATHTVNTQPRIDLT- 594
L1 1737 EESTHIVYTSANQGVVATKVTIKRATGL-----IFGSKPTIYS-KRCKRITTSVF 1792
G2 595 ---INNGNSYSFVSPDSSVALPGYMMI FVNNASV--PSVASTIIVTQ 639
L1 1793 EKDNNGRGSGSTKTKTQVQITIPA-----NGACACAMANTANTSVTK 1835

RESULT 14
: Sequence 4, Application us/06685467
: Patent No. 6060059
: GENERAL INFORMATION:
: APPLICANT: St. Gene III, Joseph W.
: APPLICANT: Barenkamp, Stephen J.
: TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flehr, Holbach, Test, Albrition & Helbert
: STREET: Four Embarradero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: United States
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in progress #10, Version #1.30
: CHILDREN APPLICATION DATA:
: AFI1 AFI-N NUMBER: us/06685467
: FILING DATE: 22-JUL-1996
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: AFI-CAL134 NUMBER: US 08/409,995
: FILING DATE: 24-MAR-1995
: ATTORNEY/AGENT INFORMATION:

```

NAME: Silva, Robin M.
REGISTRATION NUMBER: 48,304
REFERENCE: PTT Message A-61054-2/PTT-988
TELEPHONE: (415) 781-1989
TELEFAX: (415) 498-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1912 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
POLARITY: unknown
MOLECULE TYPE: protein
MISCELL: 465-467-4

Query Match: 3.7%, Score 125.5, DB 3, Length 1912
Host Local Similarity 19.1%, Pred. No. 0.15,
Mis. 157, Cons. at 170, Bt, Mismatches 276, Indels 267, Gaps 34:
6 AISCNNNAVTTGAGGTTT - - - - - PPKATPAPPTTWT - - - - - PPG 44
1195 ASNNSTLSYKLSIAKREPTTHKSALETPTTQNTAHQVQVPEHAAVYANAEVPG 1254
45 AAG - - - - - KTFHYTHAKTQVNTS - - - - - MPEKQKQNSWISHEVY 88
1255 KNAIVSAKTIIRKRIHTVTLVALEKVAEPELEKTTQTELEKVTNTPNN - - - - - L 1305
89 LSSQITWGSIVASGSMFA - - - - - DSTTKYSNRETPRARIYKVAITEANGOPTSAELNV 145
1406 ELIVAPK GANVAGHNAVTTDAITAG - - - - - TNNAR - - - - - 1339
146 FLASSYAPQGLGEMPTTPIPIVPAALPTSGPVAMSSVYNDPAFGSPGTTTS 205
1440 - - - - - GKVVVNS - - - - - NMAIAR 1354
206 SMDUSGIVSPVIVIRKHMECPGISMDNGQIV - - - - - TGSNDARKIS - - - - - 251
1455 IDPEFVAVTVAKAINTATP - - - - - VVEVNTGSAITTEPSTTQNTAMIKAVTTTKR 1411
252 - - - - - LYDSSDSNMDPQDMQVARGVQSSATMSPPRPTFGSSGSGVPPKGEVYSPSS 306
1412 KNPVPEECNTHALANDSV - - - - - KSAIVSPHSLQVCKAVNISTELKCLMANKS 1466
102 KIV - - - - - TSLPNAVNMLTADKATVPSNNAHMFGRKKS - - - - - VEQAG- 349
1467 KIRGAMNHNELASTTHILLNSGATINELNGITLNL - - - - - KKAASVKNVILNACW 1520
450 - - - - - ISIAMNNY - - - - - YTSQGLVKSAGKQSNFGVAPDAGG - - - - - NAVM 388
1521 NMGVAFASANNVINDIVATYDVEVSSREKTTTSVIVISKDQKTEVKTGAKTSVT 1580
489 YLAVKCKTTFEGSDPDQSDATINAHITLGGPQTSNPTAFASGLPFAKTFHYSV - - - - - 445
1581 KNPVPEECNTHALANDSV - - - - - GAVIVIPDCKNPEGCLVIAKAVLDAVKA 1628
446 - - - - - VLPDST - - - - - FTGPGRGSIPPEDSIPVPELVEQVPEVQNDPSLVEVHS 497
1629 GNVKVTIGAMQNDHFAIVASCTVTFADNGT - - - - - TAAQSLIVKY - - - - - 1677
496 - - - - - LKLLVQKVRKAGGCTGCTTHHTAGCTTN - - - - - YLRSGNLAIRKRIETSI 551
1678 - - - - - NKKVAVGELKIKRIKIVADTIVLVADGKVATPANNQXGKRPXDASGLAGLCKLSXTATA 1736
552 - - - - - GSVKVGKRTITSDSSISKAS - - - - - IPVATHTVMTDQPIPTIT - - - - - 594
1747 GKHGFEIVPANKAGGZEVKAKAKVAFKAGNNNA - - - - - IKSSKQETYSLKCLKLKLTSVE 1792
595 - - - - - TNNQNSYSPAVPSTGVVALPYWMLFVNNSAV - - - - - PSVASTPTPTQ 639
1794 FRIANATGGBSRIKIKEDGILITPA - - - - - NGAGAGANIANITSVTK 1835

RESULT 15
US-09-894-818B-3
Sequence 3, Application US/08894818B
Patent No. 6,261,822
GENERAL INFORMATION:
APPLICANT: TAKAKURA, HIKARU
APPLICANT: MORISHITA, MIO
APPLICANT: YAMAKOTO, Katsuhiko
APPLICANT: KITTA, Masanori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Ikunoshi
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESS: Broadway and Newmark
STREET: 419 Seventh Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
CITE: 200904
COMPUTER READABLE FORM:
MEDIA TYPE: floppy disk
COMPUTER: IBM PC compatible
DEFAULTING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,818B
FILING DATE: 20-MAY-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03253
FILING DATE: 07-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 333285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Broadway, Roger L.
PREPARATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA-1
TELEPHONE: (202) 628-5197
FAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 522 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURES:
OTHER INFORMATION: /note Xaa at position 428 is Gly or Val.
US-08-894-818B-3
Query Match: 3.6%, Score 125, DB 4, Length 522
Host Local Similarity 18.9%, Pred. No. 0.023,
Matches 89, Conservative 58, Mismatches 151, Indels 172, Gaps 20:
50 SSGIMWGSIVASGSMFASTTRYSNFTTPAPVYVPLVATTFANQGVWISAEINVDAS 149
140 SSGITDALSGAVNAAMDAITVVVAAGNSGNRY - - - - - TTGSPAAASKVTVVAVD 191
150 SYTAPQPGIAPGMPPTTDLPIVP - - - - - AAAATFTSGVIMSSVYNNAPFGSPG 199
192 KIV - - - - - SPSSK - - - - - GVALGKIKITVVAAGHMLAAAS - - - - - SMAG 1 - - - - - NATYIAPG 244
200 GTTIT - - - - - SSMDP - - - - - STGIVSDRIVYIKHMEG - - - - - PG 230
245 ISMATPHVAGIALLLQAHHSWTFDKVIML - - - - - ETADIVKPREIADIVAGACRVNAYKA 302
231 ISMDNGQIVVVG - - - - - GNDARKKRSIYSSSSSWIPGPIPMYVAVPGVQSSATGPPVPTTTC 288

```
DB 303 INEDNARKVITGVANKKQJHQLVTSQNSFV-----TAIL----- 339
QY 289 SWSTVPPNTHVSTSEPTWSTLNKYMMLTA PROLYNSQNHAWLQWAKGS 344
DB 340 YMDNANSMLDELYDPKRN ... QVYSTAYVGFERKQVYNNPLDQWTL----- 385
QY 445 VFLAGPSTAMNMYTSSSTVKSACRQSDRENVAFDAMQMAVWYDAVKKILTFGGSPD 404
DB 386 -----KVSYSGSAN 395
QY 405 YODSDATINAHILITSEPGISPN-----TVTASNSIYEARTFHTSVLPPDGSTFI 454
DB 396 YQ-VGVVSGS- SHSPGSSPSEPGPEPTVARTFNSTHYVDPSTFTNTVNSGAIKI 451
QY 455 IGOORQCIPEHSTPVYTHHVVFEQDQFTN QPNSTIVRVYHSISILLP 503
DB 452 TN ... FVPTQVYH ... LLYLLENJLVCESEFTNO YEHVEYLLP 493
```

Search completed: January 3, 2002, 10:02:47
Job time: 127 Sec



GenCode version 4.5
Copyright 1997-2000 Regeneron

protein search using sw model

Run on: January 3, 2002 05:24:38 Search time 75.07 seconds
(without alignments)
623,046 Million cell operations/sec

File: 08-06-722-602-15
Protein score: 4476
Sequence: 1 AASAPLPSATSRGMAVAVTSS 117
117MNSNVPSSVSTLEVTG 533

Scoring tables:

BLASTN62
Gap: 16.00 Gap: 0.5

Search: 52747 gaps 743720 potential hits

Minimum DB seq length: 0
Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Ageneset-1101*

```
1: 217273 1173 protein Homo H. AAI987.DAT*
2: 217273 1173 protein Homo H. AAI987.DAT*
3: 217273 1173 protein Homo H. AAI987.DAT*
4: 217273 1173 protein Homo H. AAI987.DAT*
5: 217273 1173 protein Homo H. AAI987.DAT*
6: 217273 1173 protein Homo H. AAI987.DAT*
7: 217273 1173 protein Homo H. AAI987.DAT*
8: 217273 1173 protein Homo H. AAI987.DAT*
9: 217273 1173 protein Homo H. AAI987.DAT*
10: 217273 1173 protein Homo H. AAI987.DAT*
11: 217273 1173 protein Homo H. AAI987.DAT*
12: 217273 1173 protein Homo H. AAI987.DAT*
13: 217273 1173 protein Homo H. AAI987.DAT*
14: 217273 1173 protein Homo H. AAI987.DAT*
15: 217273 1173 protein Homo H. AAI987.DAT*
16: 217273 1173 protein Homo H. AAI987.DAT*
17: 217273 1173 protein Homo H. AAI987.DAT*
18: 217273 1173 protein Homo H. AAI987.DAT*
19: 217273 1173 protein Homo H. AAI987.DAT*
20: 217273 1173 protein Homo H. AAI987.DAT*
21: 217273 1173 protein Homo H. AAI987.DAT*
22: 217273 1173 protein Homo H. AAI987.DAT*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	DB	Protein
1	3416	99.7	699	22	AA03582	Amino acid sequence
2	3416	99.7	681	22	AA03582	Synthetic amino acid
3	2217	61.7	679	21	AA03582	Fusarium venenatum
4	292	8.5	594	21	AA03582	Arabidopsis thaliana
5	292	8.5	594	21	AA03582	Arabidopsis thaliana
6	292	8.5	594	21	AA03582	Arabidopsis thaliana
7	219.5	6.4	223	21	AA03582	Arabidopsis thaliana
8	219.5	6.4	223	21	AA03582	Arabidopsis thaliana
9	219.5	6.4	223	21	AA03582	Arabidopsis thaliana
10	147	4.3	1115	20	AA03582	Staphylococcus aureus
11	144	4.3	1115	20	AA03582	Staphylococcus aureus

12	142.5	4.2	521	22	AA03582	Human colon cancer
13	142.5	4.2	521	22	AA03582	Human protein sequ
14	142.5	4.2	521	22	AA03582	Human protein sequ
15	141.5	4.1	632	20	AA03582	Human protein sequ
16	139.5	4.1	436	22	AA03582	Human protein sequ
17	139	4.1	768	22	AA03582	Human protein sequ
18	138.5	4.0	2383	22	AA03582	Human protein sequ
19	138.5	4.0	2870	21	AA03582	Human protein sequ
20	138.5	4.0	3178	21	AA03582	Human protein sequ
21	137.5	4.0	957	21	AA03582	Human protein sequ
22	137.5	4.0	957	21	AA03582	Human protein sequ
23	137	4.0	236	21	AA03582	Human protein sequ
24	137	4.0	236	21	AA03582	Human protein sequ
25	136	4.0	424	22	AA03582	Human protein sequ
26	135.5	4.0	1457	20	AA03582	Human protein sequ
27	135	3.9	441	21	AA03582	Human protein sequ
28	135	3.9	631	22	AA03582	Human protein sequ
29	134.5	3.9	851	21	AA03582	Human protein sequ
30	134	3.9	441	21	AA03582	Human protein sequ
31	134	3.9	619	22	AA03582	Human protein sequ
32	133.5	3.9	641	21	AA03582	Human protein sequ
33	132.5	3.9	1186	22	AA03582	Human protein sequ
34	132.5	3.9	1457	21	AA03582	Human protein sequ
35	132.5	3.9	1457	21	AA03582	Human protein sequ
36	132.5	3.9	1457	21	AA03582	Human protein sequ
37	131.5	3.8	330	21	AA03582	Human protein sequ
38	131.5	3.8	336	21	AA03582	Human protein sequ
39	131.5	3.8	418	21	AA03582	Human protein sequ
40	131.5	3.8	1548	20	AA03582	Human protein sequ
41	131.5	3.8	1548	20	AA03582	Human protein sequ
42	131.5	3.8	1548	20	AA03582	Human protein sequ
43	131	3.8	1548	20	AA03582	Human protein sequ
44	131	3.8	1548	20	AA03582	Human protein sequ
45	130.5	3.8	1548	20	AA03582	Human protein sequ

ALIGNMENTS

RESULT	1
AA03582	AA03582 standard, protein, 680 AA.
AA03582	AA03582 (first entry)
AA03582	Amino acid sequence of a galactose oxidase.
AA03582	Arabinoside gamma-lyase; mada gene, free folding energy, gene shuffling, directed evolution; molecular breeding, galactose oxidase.
AA03582	Hypomyces rosellus.
AA03582	W020015342-A2.
AA03582	02-AUG-2001.
AA03582	31-JAN-2001: 2001MO-0503186.
AA03582	21-JAN-2001: 2001MO-049921.
AA03582	08-DEC-2001: 2001MO-074237.
AA03582	(HIOC-) BIOCATALYTICS INC.
AA03582	Kozaki DJ, Bai P, Hua L.
AA03582	W01-18335/52.
AA03582	N-PSDB: AA03582.
AA03582	Designing synthetic nucleic acid sequences for improved amplification, expression in host cell, by comparing free energy of folding of a starting polynucleotide and a modified polynucleotide having a codon


```
XX 17- JUN-2000 (first entry)
XX
XX At 14.4 pairs Abolition protein fragment SEQ ID NO: 25862.
XX
XX 17- JUN-2000 Signal Transduction pathway; not abolic pathway;
XX hybridisation assay; spotted mapping; gene expression control; promoter;
XX hybridisation sequence.
XX
XX Available at:
XX
XX EMBL: U44015.AZ.
XX
XX 06- SEP-2000.
XX
XX 17- FEB-2000 Zm0001P.040143.9.
XX
XX 25- FEB-1999 9908-0121825.
XX 05- MAR-1999 9908-0123180.
XX 09- MAR-1999 9908-0123548.
XX 23- MAR-1999 9908-0125788.
XX 25- MAR-1999 9908-0126264.
XX 25- MAR-1999 9908-0126785.
XX 01- APR-1999 9908-0127462.
XX 04- APR-1999 9908-0128234.
XX 08- APR-1999 9908-0128714.
XX 11- APR-1999 9908-0129845.
XX 14- APR-1999 9908-0130077.
XX 21- APR-1999 9908-0130849.
XX 28- APR-1999 9908-0130810.
XX 28- APR-1999 9908-0130891.
XX 28- APR-1999 9908-0131449.
XX 30- APR-1999 9908-0132048.
XX 03- APR-1999 9908-0132407.
XX 04- MAY-1999 9908-0132484.
XX 05- MAY-1999 9908-0132485.
XX 06- MAY-1999 9908-0132486.
XX 06- MAY-1999 9908-0132487.
XX 07- MAY-1999 9908-0132863.
XX 11- MAY-1999 9908-0134256.
XX 14- MAY-1999 9908-0134218.
XX 14- MAY-1999 9908-0134219.
XX 14- MAY-1999 9908-0134321.
XX 15- MAY-1999 9908-0134321.
XX 15- MAY-1999 9908-0134370.
XX 15- MAY-1999 9908-0134378.
XX 19- MAY-1999 9908-0134941.
XX 20- MAY-1999 9908-0135124.
XX 21- MAY-1999 9908-0135353.
XX 24- MAY-1999 9908-0135629.
XX 25- MAY-1999 9908-0136021.
XX 27- MAY-1999 9908-0136392.
XX 28- MAY-1999 9908-0137282.
XX 01- JUN-1999 9908-0137222.
XX 03- JUN-1999 9908-0137528.
XX 04- JUN-1999 9908-0137528.
XX 07- JUN-1999 9908-0137724.
XX 08- JUN-1999 9908-0138094.
XX 08- JUN-1999 9908-0138540.
XX 10- JUN-1999 9908-0138847.
XX 14- JUN-1999 9908-0139119.
XX 14- JUN-1999 9908-0139452.
XX 14- JUN-1999 9908-0139453.
XX 17- JUN-1999 9908-0139492.
XX 18- JUN-1999 9908-0139454.
XX 18- JUN-1999 9908-0139455.
XX 18- JUN-1999 9908-0139456.
XX 18- JUN-1999 9908-0139457.
XX 18- JUN-1999 9908-0139458.
XX 18- JUN-1999 9908-0139459.
XX 18- JUN-1999 9908-0139460.
XX 18- JUN-1999 9908-0139461.
XX 18- JUN-1999 9908-0139462.
XX 18- JUN-1999 9908-0139463.
XX
XX 18- JUN-1999 9908-0139750.
XX 18- JUN-1999 9908-0139763.
XX 21- JUN-1999 9908-0139817.
XX 22- JUN-1999 9908-0139899.
XX 23- JUN-1999 9908-0140153.
XX 23- JUN-1999 9908-0140354.
XX 24- JUN-1999 9908-0140695.
XX 28- JUN-1999 9908-0140823.
XX 29- JUN-1999 9908-0140991.
XX 30- JUN-1999 9908-0141287.
XX 01- JUL-1999 9908-0141842.
XX 01- JUL-1999 9908-0142154.
XX 02- JUL-1999 9908-0142055.
XX 06- JUL-1999 9908-0142390.
XX 08- JUL-1999 9908-0142803.
XX 09- JUL-1999 9908-0142920.
XX 12- JUL-1999 9908-0142972.
XX 13- JUL-1999 9908-0143942.
XX 14- JUL-1999 9908-0143624.
XX 15- JUL-1999 9908-0144005.
XX 16- JUL-1999 9908-0144085.
XX 16- JUL-1999 9908-0144086.
XX 19- JUL-1999 9908-0144325.
XX 19- JUL-1999 9908-0144331.
XX 19- JUL-1999 9908-0144332.
XX 19- JUL-1999 9908-0144333.
XX 19- JUL-1999 9908-0144334.
XX 19- JUL-1999 9908-0144335.
XX 20- JUL-1999 9908-0144352.
XX 20- JUL-1999 9908-0144632.
XX 20- JUL-1999 9908-0144884.
XX 21- JUL-1999 9908-0144814.
XX 21- JUL-1999 9908-0145086.
XX 21- JUL-1999 9908-0145088.
XX 22- JUL-1999 9908-0145085.
XX 22- JUL-1999 9908-0145087.
XX 22- JUL-1999 9908-0145089.
XX 22- JUL-1999 9908-0145192.
XX 22- JUL-1999 9908-0145145.
XX 23- JUL-1999 9908-0145218.
XX 23- JUL-1999 9908-0145224.
XX 24- JUL-1999 9908-0145276.
XX 26- JUL-1999 9908-0145913.
XX 27- JUL-1999 9908-0145918.
XX 27- JUL-1999 9908-0145919.
XX 28- JUL-1999 9908-0145951.
XX 02- AUG-1999 9908-0146386.
XX 02- AUG-1999 9908-0146388.
XX 02- AUG-1999 9908-0146389.
XX 03- AUG-1999 9908-0147038.
XX 04- AUG-1999 9908-0147204.
XX 04- AUG-1999 9908-0147302.
XX 05- AUG-1999 9908-0147192.
XX 05- AUG-1999 9908-0147260.
XX 06- AUG-1999 9908-0147403.
XX 06- AUG-1999 9908-0147416.
XX 06- AUG-1999 9908-0147493.
XX 09- AUG-1999 9908-0147935.
XX 10- AUG-1999 9908-0148171.
XX 11- AUG-1999 9908-0148319.
XX 12- AUG-1999 9908-0148341.
XX 13- AUG-1999 9908-0148565.
XX 13- AUG-1999 9908-0148684.
XX 16- AUG-1999 9908-0149458.
XX 17- AUG-1999 9908-0149175.
XX 18- AUG-1999 9908-0149426.
XX 20- AUG-1999 9908-0149222.
XX 20- AUG-1999 9908-0149723.
XX 20- AUG-1999 9908-0149929.
XX 23- AUG-1999 9908-0149902.
XX 23- AUG-1999 9908-0149930.
XX 25- AUG-1999 9908-0150566.
XX 26- AUG-1999 9908-0150884.
```


11 MAY 1999 9905-0134219. PR 23-JUL-1999 9905-0145224.
14 MAY 1999 9905-0134219. PR 26-JUL-1999 9905-0145226.
14 MAY 1999 9905-0134221. PR 27-JUL-1999 9905-0145913.
14 MAY 1999 9905-0134270. PR 27-JUL-1999 9905-0145918.
18 MAY 1999 9905-0134768. PR 27-JUL-1999 9905-0145919.
19 MAY 1999 9905-0134941. PR 28-JUL-1999 9905-0145951.
20 MAY 1999 9905-0135124. PR 02-AUG-1999 9905-0146386.
21 MAY 1999 9905-0135223. PR 02-AUG-1999 9905-0146388.
24 MAY 1999 9905-0135629. PR 02-AUG-1999 9905-0146389.
26 MAY 1999 9905-0136021. PR 03-AUG-1999 9905-0147038.
27 MAY 1999 9905-0136392. PR 04-AUG-1999 9905-0147204.
28 MAY 1999 9905-0136722. PR 04-AUG-1999 9905-0147302.
01-JUN-1999 9905-0137222. PR 05-AUG-1999 9905-0147192.
03-JUN-1999 9905-0137628. PR 05-AUG-1999 9905-0147260.
04-JUN-1999 9905-0137502. PR 05-AUG-1999 9905-0147203.
07-JUN-1999 9905-0137724. PR 06-AUG-1999 9905-0147416.
08-JUN-1999 9905-0138094. PR 09-AUG-1999 9905-0147493.
10-JUN-1999 9905-0138540. PR 09-AUG-1999 9905-0147935.
11-JUN-1999 9905-0138847. PR 10-AUG-1999 9905-0148171.
14-JUN-1999 9905-0139119. PR 11-AUG-1999 9905-0148319.
14-JUN-1999 9905-0139452. PR 12-AUG-1999 9905-0148341.
16-JUN-1999 9905-0139453. PR 13-AUG-1999 9905-0148565.
17-JUN-1999 9905-0139492. PR 13-AUG-1999 9905-0148684.
18-JUN-1999 9905-0139454. PR 16-AUG-1999 9905-0149368.
18-JUN-1999 9905-0139455. PR 17-AUG-1999 9905-0149175.
18-JUN-1999 9905-0139456. PR 18-AUG-1999 9905-0149426.
18-JUN-1999 9905-0139457. PR 20-AUG-1999 9905-0149722.
18-JUN-1999 9905-0139458. PR 20-AUG-1999 9905-0149723.
18-JUN-1999 9905-0139459. PR 20-AUG-1999 9905-0149929.
18-JUN-1999 9905-0139460. PR 23-AUG-1999 9905-0149902.
18-JUN-1999 9905-0139461. PR 23-AUG-1999 9905-0149930.
18-JUN-1999 9905-0139462. PR 25-AUG-1999 9905-0150566.
18-JUN-1999 9905-0139463. PR 26-AUG-1999 9905-0150884.
18-JUN-1999 9905-0139750. PR 27-AUG-1999 9905-0151065.
18-JUN-1999 9905-0139763. PR 27-AUG-1999 9905-0151066.
21-JUN-1999 9905-0139817. PR 27-AUG-1999 9905-0151080.
22-JUN-1999 9905-0139847. PR 30-AUG-1999 9905-0151303.
23-JUN-1999 9905-0140353. PR 31-AUG-1999 9905-0151348.
24-JUN-1999 9905-0140354. PR 01-SEP-1999 9905-0151930.
24-JUN-1999 9905-0140605. PR 07-SEP-1999 9905-0152363.
28-JUN-1999 9905-0140823. PR 10-SEP-1999 9905-0153070.
29-JUN-1999 9905-0140901. PR 13-SEP-1999 9905-0153758.
30-JUN-1999 9905-0141287. PR 15-SEP-1999 9905-0154018.
01-JUL-1999 9905-0141842. PR 16-SEP-1999 9905-0154039.
01-JUL-1999 9905-0142154. PR 20-SEP-1999 9905-0154779.
02-JUL-1999 9905-0142055. PR 22-SEP-1999 9905-0155139.
06-JUL-1999 9905-0142390. PR 23-SEP-1999 9905-0155486.
08-JUL-1999 9905-0142803. PR 24-SEP-1999 9905-0155659.
09-JUL-1999 9905-0142920. PR 28-SEP-1999 9905-0156458.
12-JUL-1999 9905-0142977. PR 29-SEP-1999 9905-0156596.
13-JUL-1999 9905-0143542. PR 04-OCT-1999 9905-0157117.
14-JUL-1999 9905-0143624. PR 05-OCT-1999 9905-0157753.
15-JUL-1999 9905-0144005. PR 06-OCT-1999 9905-0157865.
16-JUL-1999 9905-0144085. PR 07-OCT-1999 9905-0158029.
16-JUL-1999 9905-0144086. PR 08-OCT-1999 9905-0158232.
19-JUL-1999 9905-0144332. PR 12-OCT-1999 9905-0158369.
19-JUL-1999 9905-0144331. PR 13-OCT-1999 9905-0159293.
19-JUL-1999 9905-0144332. PR 13-OCT-1999 9905-0159294.
19-JUL-1999 9905-0144333. PR 13-OCT-1999 9905-0159295.
19-JUL-1999 9905-0144334. PR 14-OCT-1999 9905-0159429.
19-JUL-1999 9905-0144335. PR 14-OCT-1999 9905-0159530.
20-JUL-1999 9905-0144352. PR 14-OCT-1999 9905-0159531.
20-JUL-1999 9905-0144353. PR 14-OCT-1999 9905-0159637.
20-JUL-1999 9905-0144354. PR 18-OCT-1999 9905-0159638.
21-JUL-1999 9905-0144384. PR 18-OCT-1999 9905-0159584.
21-JUL-1999 9905-0144385. PR 21-OCT-1999 9905-0160741.
21-JUL-1999 9905-0144386. PR 21-OCT-1999 9905-0160741.
21-JUL-1999 9905-0144387. PR 21-OCT-1999 9905-0160767.
22-JUL-1999 9905-0145085. PR 21-OCT-1999 9905-0160768.
22-JUL-1999 9905-0145087. PR 21-OCT-1999 9905-0160770.
22-JUL-1999 9905-0145089. PR 21-OCT-1999 9905-0160814.
22-JUL-1999 9905-0145102. PR 21-OCT-1999 9905-0160815.
23-JUL-1999 9905-0145145. PR 22-OCT-1999 9905-0160980.
23-JUL-1999 9905-0145218. PR 22-OCT-1999 9905-0160981.

26-JUN-1999	990S-0140823	PR	10-SEP-1999	990S-0153070
26-JUN-1999	990S-0140991	PR	13-SEP-1999	990S-0153758
30-JUN-1999	990S-0141287	PR	15-SEP-1999	990S-0154018
01-JUL-1999	990S-0141842	PR	16-SEP-1999	990S-0154039
01-JUL-1999	990S-0142054	PR	20-SEP-1999	990S-0154779
02-JUL-1999	990S-0142055	PR	22-SEP-1999	990S-0155139
06-JUL-1999	990S-0142330	PR	23-SEP-1999	990S-0156486
08-JUL-1999	990S-0142803	PR	24-SEP-1999	990S-0156564
09-JUL-1999	990S-0142920	PR	28-SEP-1999	990S-0156458
12-JUL-1999	990S-0142377	PR	29-SEP-1999	990S-0156506
13-JUL-1999	990S-0143542	PR	04-OCT-1999	990S-0157117
14-JUL-1999	990S-0143624	PR	05-OCT-1999	990S-0157753
15-JUL-1999	990S-0144005	PR	06-OCT-1999	990S-0157865
16-JUL-1999	990S-0144085	PR	07-OCT-1999	990S-0158029
19-JUL-1999	990S-0144332	PR	08-OCT-1999	990S-0158232
19-JUL-1999	990S-0144333	PR	12-OCT-1999	990S-0158369
19-JUL-1999	990S-0144334	PR	13-OCT-1999	990S-0159293
20-JUL-1999	990S-0144352	PR	13-OCT-1999	990S-0159294
20-JUL-1999	990S-0144884	PR	13-OCT-1999	990S-0159295
21-JUL-1999	990S-0144814	PR	14-OCT-1999	990S-0159329
21-JUL-1999	990S-0145086	PR	14-OCT-1999	990S-0159330
21-JUL-1999	990S-0145085	PR	14-OCT-1999	990S-0159331
22-JUL-1999	990S-0145087	PR	14-OCT-1999	990S-0159637
22-JUL-1999	990S-0145089	PR	18-OCT-1999	990S-0159638
22-JUL-1999	990S-0145145	PR	18-OCT-1999	990S-0159684
23-JUL-1999	990S-0145218	PR	22-OCT-1999	990S-0160741
26-JUL-1999	990S-0145276	PR	21-OCT-1999	990S-0160767
27-JUL-1999	990S-0145913	PR	21-OCT-1999	990S-0160770
27-JUL-1999	990S-0145919	PR	21-OCT-1999	990S-0160778
28-JUL-1999	990S-0145951	PR	21-OCT-1999	990S-0160814
02-AUG-1999	990S-0146386	PR	21-OCT-1999	990S-0160815
02-AUG-1999	990S-0146388	PR	22-OCT-1999	990S-0160980
02-AUG-1999	990S-0146389	PR	22-OCT-1999	990S-0160981
03-AUG-1999	990S-0147038	PR	22-OCT-1999	990S-0160989
04-AUG-1999	990S-0147204	PR	25-OCT-1999	990S-0161404
04-AUG-1999	990S-0147302	PR	25-OCT-1999	990S-0161405
05-AUG-1999	990S-0147192	PR	27-OCT-1999	990S-0161406
05-AUG-1999	990S-0147265	PR	26-OCT-1999	990S-0161359
06-AUG-1999	990S-0147303	PR	26-OCT-1999	990S-0161360
06-AUG-1999	990S-0147416	PR	26-OCT-1999	990S-0161361
09-AUG-1999	990S-0147443	PR	28-OCT-1999	990S-0161920
09-AUG-1999	990S-0147935	PR	28-OCT-1999	990S-0161993
10-AUG-1999	990S-0148171	PR	29-OCT-1999	990S-0162142
11-AUG-1999	990S-0148319	PR		
12-AUG-1999	990S-0148341	PR		
13-AUG-1999	990S-0148565	PR		
16-AUG-1999	990S-0149368	PR		
17-AUG-1999	990S-0149375	PR		
18-AUG-1999	990S-0149426	PR		
20-AUG-1999	990S-0149722	PR		
20-AUG-1999	990S-0149723	PR		
23-AUG-1999	990S-0149982	PR		
23-AUG-1999	990S-0149983	PR		
25-AUG-1999	990S-0150566	PR		
26-AUG-1999	990S-0150684	PR		
27-AUG-1999	990S-0151065	PR		
27-AUG-1999	990S-0151066	PR		
27-AUG-1999	990S-0151080	PR		
30-AUG-1999	990S-0151303	PR		
31-AUG-1999	990S-0151438	PR		
01-SEP-1999	990S-0151530	PR		
07-SEP-1999	990S-0152363	PR		

Query Match: 8.48; Score 286.5; DB 21; Length 504;

Best Local Similarity: 25.28; Pred. No. 3.8e-13;

Matches 124; Conservative 62; Mismatches 203; Indels 103; Gaps 23;

QY	208	DPECTGVSDFVTIVTKHDMCEPGISMGNGIVTGG-NDAKTISLYSSSDS--WIRGP	264
DB	53	dvntgalkfqlst..dtwssqjlltvglltystjg7g7gntnrylssconvwvwyf	109
QY	265	DMQVA-WYQSSATMSPCRVTTTGSMSGVFFKNGEYSSSKTWTSLPNKXVPMILIA	323
DB	110	galadkriwyslqallpdkkltivlqtr..	154
QY	324	DKQCIYRSD---NHAWFCW-----KRGSV---PQACPSTAMNW	356
DB	155	dsl rrtdfpeennl pvlvlnldgullllamnsllspktaqyikelp ppqarr-	213
QY	357	YYTSGSTNWKSAQKPNQSNWYAPLPMQSNNAVWYTAVKSKIIITFGSPRYQSDATTNHII	416
DB	214	ypasgsallfpjlylvknpkvlp.....acvtyvqgask qdyzykagpki	257
QY	417	TTLEERC-----TSPNVVPASNCIYFAVPTHTSVV pncstf TTCGGPFGIP--FE	465
DB	258	g..cpalqdcailinsakp..kltcmqf..f..mslcczlpogd nqaktrqcsjwyf	314
QY	466	DSIVVF P P YVPRQ---TFYKQNPNSIVPAYSISLLEKGVNPNQOQOQOCTIN	521
DB	315	gkdpdapl kplaaagktrtdq..kplclpmpyhsas fjfkltvqgsrfrdqygvw	374

[illegible]

PR	04-AUG-1999	9905-01472704
PR	04-AUG-1999	9905-01473702
PR	05-AUG-1999	9905-01471922
PR	05-AUG-1999	9905-01472600
PR	05-AUG-1999	9905-01473703
PR	06-AUG-1999	9905-01472716
PR	06-AUG-1999	9905-01472454
PR	09-AUG-1999	9905-01479335
PR	10-AUG-1999	9905-01481871
PR	12-AUG-1999	9905-01483819
PR	13-AUG-1999	9905-01483841
PR	13-AUG-1999	9905-01485655
PR	14-AUG-1999	9905-01485684
PR	14-AUG-1999	9905-01484568
PR	18-AUG-1999	9905-01491775
PR	18-AUG-1999	9905-01494265
PR	20-AUG-1999	9905-01491922
PR	20-AUG-1999	9905-01492723
PR	20-AUG-1999	9905-01492729
PR	23-AUG-1999	9905-01493902
PR	23-AUG-1999	9905-01499330
PR	25-AUG-1999	9905-01505566
PR	26-AUG-1999	9905-01508884
PR	27-AUG-1999	9905-01510655
PR	27-AUG-1999	9905-01510666
PR	27-AUG-1999	9905-01510800
PR	30-AUG-1999	9905-01513303
PR	31-AUG-1999	9905-01514338
PR	01-SEP-1999	9905-01519330
PR	01-SEP-1999	9905-01523263
PR	01-SEP-1999	9905-01530770
PR	13-SEP-1999	9905-01537558
PR	15-SEP-1999	9905-01540418
PR	16-SEP-1999	9905-01545039
PR	20-SEP-1999	9905-01545479
PR	22-SEP-1999	9905-01551539
PR	23-SEP-1999	9905-01554886
PR	24-SEP-1999	9905-01555659
PR	28-SEP-1999	9905-01564588
PR	29-SEP-1999	9905-01565696
PR	04-OCT-1999	9905-01571717
PR	05-OCT-1999	9905-01577533
PR	06-OCT-1999	9905-01578655
PR	07-OCT-1999	9905-01580229
PR	08-OCT-1999	9905-01582632
PR	12-OCT-1999	9905-01584366
PR	13-OCT-1999	9905-01592933
PR	13-OCT-1999	9905-01592954
PR	13-OCT-1999	9905-01593329
PR	14-OCT-1999	9905-01593330
PR	14-OCT-1999	9905-01593331
PR	14-OCT-1999	9905-01596637
PR	14-OCT-1999	9905-01596638
PR	18-OCT-1999	9905-01595884
PR	21-OCT-1999	9905-01607241
PR	21-OCT-1999	9905-01607247
PR	21-OCT-1999	9905-01607277
PR	21-OCT-1999	9905-01607768
PR	21-OCT-1999	9905-01607770
PR	21-OCT-1999	9905-01614055
PR	25-OCT-1999	9905-01614065
PR	26-OCT-1999	9905-01613359
PR	26-OCT-1999	9905-01613650
PR	26-OCT-1999	9905-01613661
PR	28-OCT-1999	9905-01613661
PR	28-OCT-1999	9905-01616200
PR	28-OCT-1999	9905-01619933
PR	28-OCT-1999	9905-01619933

1041 T. Iseki, T. Nishikawa, T. Hayashi, K. Saito, K. Yamamoto, J.
Ishii, S. Sudo, Y. Watanabe, A. Hagi, Y. Otsu, T.
W. 2001 48739/44.

Primer sets for synthesizing polynucleotides, particularly the 5602
full-length cDNAs defined in the specification, and for the detection
and/or diagnosis of the abnormality of the proteins encoded by the
full-length cDNAs -

Claim 8: SEQ ID 12047; 2537bp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602
full-length cDNAs defined in the specification. Where a primer set
comprises: (a) an oligo dT primer and an oligonucleotide complementary
to the complementary strand of a polynucleotide which comprises one of
the 5602 nucleotide sequences defined in the specification, where the
oligonucleotide comprises at least 15 nucleotides; or (b) a combination
of an oligonucleotide comprising a sequence complementary to the
complementary strand of a polynucleotide which comprises a 5' end
sequence and an oligonucleotide comprising a sequence complementary to a
polynucleotide which comprises a 3' end sequence, where the
oligonucleotide comprises at least 15 nucleotides and the combination of
the 5' end sequence and the 3' end sequence is selected from those defined in
the specification. The primer sets can be used in antisense therapy and
in gene therapy. The primers are useful for synthesizing polynucleotides,
particularly full length cDNAs. The primers are also useful for the
detection and/or diagnosis of the abnormality of the proteins encoded by
the full-length cDNAs. The primers allow containing of the full-length
cDNAs easily without any specialised methods. AAB67050 to AAB13628 and
AAB13629 to AAB13642 represent human cDNA sequences. AAB92316 to
AAB15680 represent human amino acid sequences and AAB13629 to AAB13632
represent oligonucleotides, all of which are used in the exemplification
of the present invention.

Sequence 600 AA:

Query Match 4.28; Score 142.5; PB 22; Length 600;
Best Local Similarity 19.14; Pred. No. 0.028;
Matches 93; Conservative 71; Mismatches 107; Indels 157; Gaps 23;

28 NAAAGNGLIIVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 87
107 LKLLIIVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 224
88 YVSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 143
225 LSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 269
144 LSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 184
270 LSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 329
165 MSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 244
400 AAGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 380
245 NAAFKLSIYVSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 298
301 LSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 349
299 GGVYSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 335
140 GGVYSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 496
400 GGVYSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 386
407 LSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 534
400 GGVYSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 446

Db 535 GGVYSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 572
C7 447 LSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 454
Db 573 LSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 578

RESULT 14
AAB67050
ID AAB67050 standard; protein, 642 AA.

XX AAB67050;

XX 09-APR-2001 (first entry)

XX Human immune response molecule (IMUN) protein seq ID NO: 4.

XX Human: IMUN; immune response molecule; autoimmune disorder;

XX inflammatory disorder; cell proliferation disorder; cancer;

XX Homo sapiens.

XX K02601:2567-A2.

XX 11-JAN 2001.

XX 06-JUL-2000; 2000MO-US18505.

XX 06-JUL-1999; 9908 014572.

XX 09-SEP-1999; 9905-0153170.

XX (INCY) INCYTE GENOMICS INC.

XX Tang YL, Yue H, Yang J, Azimzai Y, Roughton MP, Lu DM;

XX WPI: 2001-123113/13.

XX N-PSDB: AAF31247.

XX Novel human immune response molecules (IMUN), useful for diagnosis,

XX treating and preventing disorders associated with abnormal expression

XX of IMUN, e.g. Addison's disease, allergies, anemia, asthma and

XX atherosclerosis -

XX disclosure, page 83-84; 95pp; English.

XX Sequence 642 AA;

Query Match 4.28; Score 142.5; PB 22; Length 642;
Best Local Similarity 19.14; Pred. No. 0.03;
Matches 93; Conservative 69; Mismatches 170; Indels 155; Gaps 23;

28 NAAAGNGLIIVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 87
107 LKLLIIVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 226
88 YVSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 143
225 LSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 266
144 LSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 185

Db 229 LKLLIIVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 412

Db 267 LSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 412

C7 144 LSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 185

Db 313 LSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 472

C7 186 MSSGFWVWVYANRPPRIIVIMATLVNWNLSMTPEQVQNMICKEHY 245

GenCode version 4.5
Copyright (c) 1995 - 2002 Compugen Ltd.

M protein - protein search, using SW model

January 3, 2002, 09:11:13 : Search time 39.08 seconds

(without alignments)
1445,536 Million cell updates/sec

US-09-722-602-15

Perfect score:
3425
1 ASAPISALSRRNMWATVCSAAGSAGNECNKATIDGKDTFWITFGANDPKPPHYTIDMK 639

Scoring table:
BLOSUM62
Gapop 10.0 : Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum hit seq length: 0
Maximum hit seq length: 206900000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR:68:
2: PIR:1:
3: PIR:3:
4: PIR:4:

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3416	99.7	728	1 A38084	galactose oxidase
2	271	7.9	544	2 B6332	hypothetical prote
3	269	7.2	547	2 I06258	probable galactose
4	267	7.8	545	2 I45935	probable galactose
5	242.5	7.1	564	2 B86278	hypothetical prote
6	203	5.9	647	2 A45244	oxy-alpha-sialids
7	202	5.9	2458	2 A84412	hypothetical prote
8	197	5.4	1044	1 NM1158	oxy-alpha-sialids
9	186.5	5.3	1361	2 J24425	hypothetical prote
10	175	5.1	559	2 A42946	glyoxal oxidase (E
11	175	5.1	1441	2 B63887	hypothetical prote
12	166.5	4.9	5186	2 B67547	probable RTX faml
13	162.5	4.7	1433	2 I30261	chitinase (EC 3.2.
14	160.5	4.7	836	2 F00349	AVIcoelase III - As
15	160	4.7	2232	2 F01143	hypothetical prote
16	151	4.4	943	2 T35197	hypothetical prote
17	148.5	4.3	627	2 G66452	hypothetical prote
18	147	4.3	1749	2 S75198	hypothetical prote
19	146	4.3	1465	2 T15660	hypothetical prote
20	145.5	4.2	1275	2 T33469	hypothetical prote
21	144.5	4.2	4199	2 S74412	hypothetical prote
22	143	4.2	1733	2 B85557	hypothetical prote
23	143	4.2	1733	2 E73567	polymorphic membra
24	143	4.2	1732	2 G81561	polymorphic membra
25	142.5	4.2	659	2 B44802	polymorphic membra
26	142.5	4.2	1289	2 T12243	subunit of protein
27	142.5	4.2	1402	2 B44858	lactocarpin (EC 3.4
28	142	4.1	1180	2 B86719	hypothetical prote
29	141.5	4.1	1461	2 A85547	hypothetical prote

30	140.5	4.1	827	2 J04900	xanthomonapepsin (
31	139.5	4.1	713	2 B75489	hypothetical prote
32	139.5	4.1	784	2 T16599	hypothetical prote
33	139.5	4.1	1649	2 G66822	hypothetical prote
34	139	4.1	1286	2 S28634	adhesin AIDA-1 pre
35	138.5	4.0	1032	2 T14473	hypothetical prote
36	138.5	4.0	2983	2 G64962	fructose 6-phosphate
37	138	4.0	1147	2 T40856	cell polarity prot
38	136.5	4.0	1356	2 T18521	beta tubulin-ii
39	136.5	4.0	1217	2 T25994	hypothetical prote
40	136.5	4.0	1599	2 S22737	glicosyltransferase
41	135	3.9	613	2 B64786	ecdc-1,4-beta-xyla
42	135	3.9	1747	2 T21894	hypothetical prote
43	135	3.9	1805	2 T21888	hypothetical prote
44	134.5	3.9	1090	2 S59077	cellulose 1,4-beta
45	134.5	3.9	1662	2 T18540	mota protein precu

ALIGNMENTS

RESULT 1
A38084
galactose oxidase (EC 1.1.3.9) precursor [validated] - fungus (Cladobotryum dendroide
C:Species: Cladobotryum dendroides
C:Title: Cladobotryum dendroides precursor [validated] - fungus (Cladobotryum dendroide
C:Accession: A38084
C:Author: M.J. Ogel, Z.B. Stevens, C.J. Yaday, K.D.S., Keen, J.R., Knowles, P.F.
J Biol Chem. 267, 8146-8152, 1992
A:Title: Galactose oxidase of Cladobotryum dendroides. Gene cloning and sequence analysis
A:Reference number: A38084; MIM:32336025
A:Accession: A38084
A:Molecule type: DNA
A:Residues: 1-728 <MCPS>
A:Cross-references: GB:M86819
A:Note: It is uncertain whether Met 1 or Met-49 is the initiator
A:Note: puts of this sequence, including the amino end of the mature protein, were c
R:Itto, N.; Phillips, S.E.V.; Stevens, C., Ogel, Z.B.; Methelsson, M.J., Keen, J.N.; Y
Nature 350, 87-90, 1991
A:Title: Novel thioether bond revealed by a 1.7 angstrom crystal structure of galacto
A:Reference number: A46241; MIM:3163641
A:Contents: annotation
R:Itto, N.; Phillips, S.E.V.; Knowles, P.F.
submitted to the Brookhaven Protein Data Bank, September 1993
A:Reference number: A51740; PDB:1GOF
A:Contents: annotation; X-ray crystallography, 1.7 angstroms, residues 90-728
R:Itto, N.; Phillips, S.E.V.; Knowles, P.F.
submitted to the Brookhaven Protein Data Bank, September 1993
A:Reference number: A51741; PDB:1GOG
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 90 728
R:Itto, N.; Phillips, S.E.V.; Knowles, P.F.
submitted to the Brookhaven Protein Data Bank, September 1993
A:Reference number: A51742; PDB:1GHH
A:Contents: annotation; X-ray crystallography, 2.2 angstroms, residues 90-728
A:Gene: gaou
A:Genetics:
C:Superfamily: Cladobotryum dendroides galactose oxidase
C:Keywords: disulfide bond; metal binding; oxidoreductase
E:164/bmann: signal sequence #status predicted <PPO>
E:164/bmann: signal sequence #status predicted <PPO>
E:164/bmann: galactose oxidase #status experimental <MAT>
E:107-116/604-607/disulfide bonds: #status experimental
E:107-116/604-607/disulfide bonds: #status experimental
E:361-584/584-670/tyrosine (Gln Tyr) #status experimental
E:419/binding site: substrate (Arg) #status predicted

Query Match 99.7% Score 3416, DB 1; Length 728;
Best local similarity 99.7% Prod. No. 9.5e+208;
Matches 637; Conserved 1; Mismatches 1; Indels 0; Gaps 0;
1 ASAPISALSRRNMWATVCSAAGSAGNECNKATIDGKDTFWITFGANDPKPPHYTIDMK 60
|||||

[illegible]

```

16 187 LIASSTAYITNENONENRATWYSANNA-----LIPWICADLISAPRVNVAIKLP-- 239
17 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

```



SMALL: SM00231: FA580: 1.
oxidoreductase; Strand; Copper; Repeat; 3D-structure.
STRAND: 1 41 GALACTOSE OXIDASE.
PEPEAT 42 680
PEPEAT 279 421 KETCH 1.
PEPEAT 424 472 KETCH 2.
PEPEAT 492 541 KETCH 3.
DISULFID 59 68
HELIX-PTH 269 414
DISULFID 554 654
METAL 414 414
METAL 546 546
METAL 547 547
METAL 622 622
STRAND 48 49
STRAND 56 59
STRAND 63 64
STRAND 64 66
STRAND 66 66
HELIX 68 72
STRAND 76 77
STRAND 80 81
HELIX 85 87
STRAND 94 112
STRAND 120 121
STRAND 123 123
STRAND 125 131
STRAND 142 145
STRAND 153 169
STRAND 172 172
STRAND 173 174
STRAND 180 181
STRAND 183 188
STRAND 197 198
STRAND 201 207
STRAND 214 217
STRAND 219 221
STRAND 224 227
STRAND 236 237
STRAND 243 249
STRAND 250 253
STRAND 254 260
STRAND 262 263
STRAND 271 274
STRAND 276 277
STRAND 279 283
STRAND 285 286
STRAND 288 289
STRAND 291 295
STRAND 296 299
STRAND 300 303
STRAND 312 312
STRAND 315 318
STRAND 320 321
STRAND 324 327
STRAND 330 340
STRAND 332 332
STRAND 340 344
STRAND 345 348
STRAND 349 353
STRAND 354 355
STRAND 359 361
STRAND 366 368
STRAND 369 372
STRAND 373 373
STRAND 378 380
STRAND 382 383
STRAND 386 388
STRAND 394 399
STRAND 402 403
STRAND 405 411
STRAND 413 414
STRAND 415 416

FT STRAND 417 418
FT STRAND 423 423
FT TURN 424 424
FT STRAND 426 431
FT TURN 432 435
FT STRAND 436 440
FT STRAND 443 443
FT STRAND 450 450
FT STRAND 452 452
FT STRAND 455 459
FT STRAND 463 464
FT STRAND 468 471
FT TURN 473 474
FT STRAND 480 481
FT STRAND 484 487
FT STRAND 488 490
FT STRAND 493 496
FT STRAND 499 499
FT STRAND 501 501
FT TURN 504 505
FT STRAND 510 510
FT STRAND 515 518
FT TURN 519 522
FT STRAND 523 526
FT TURN 536 537
FT STRAND 542 542
FT TURN 544 545
FT STRAND 548 552
FT TURN 557 558
FT STRAND 565 570
FT STRAND 572 574
FT HELIX 575 575
FT STRAND 577 578
FT TURN 581 581
FT STRAND 587 590
FT STRAND 594 596
FT TURN 597 598
FT STRAND 600 605
FT STRAND 611 616
FT STRAND 619 621
FT TURN 622 623
FT STRAND 624 624
FT STRAND 630 633
FT STRAND 646 637
FT STRAND 642 646
FT TURN 651 653
FT STRAND 657 664
FT TURN 666 667
FT STRAND 670 670
FT STRAND 674 679
SQ SEQUENCE 680 AA: 72823 MW: 2F97561P64F44P9 CP064:

Query Match 99.78; Score 3416; DB 1; Length 680;
Best local Similarity 99.78; Pred. No. 2.2e-205;
Matches 637; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASAPIGSAISPNNAVAVTCUSASQSPNPKAIDGKDKTFHFFHYGANGDPRPHYTUMK 60
DB 42 ASAPIGSAISPNNAVAVTCUSASQSPNPKAIDGKDKTFHFFHYGANGDPRPHYTUMK 101
QY 61 TIGNVNCSMI PRDQNNQNMIGRHYVLSNGTNNKTSVVASGSWFAUSTIKYSNFFTRP 120
DB 102 TIGNVNCSMI PRDQNNQNMIGRHYVLSNGTNNKTSVVASGSWFAUSTIKYSNFFTRP 161
QY 121 APYVPIVAITFANQOPSTIAELNVQASSYTAPOPGIGRWGPTIIDLIVAAAAIEPIS 180
DB 162 APYVPIVAITFANQOPSTIAELNVQASSYTAPOPGIGRWGPTIIDLIVAAAAIEPIS 221
QY 181 GVALMSSSYRNDAGGSGPGGTTTSSMDPSTGIVSDRTVYTKHMRFGGSMGNGGIV 240
DB 222 GVALMSSSYRNDAGGSGPGGTTTSSMDPSTGIVSDRTVYTKHMRFGGSMGNGGIV 281

Database version 4.5
Copyright (c) 1998 - 2000 CompuLink Ltd.

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

us-09-722-602-15

20 160.5 4.7 866 3 074170
21 160 4.7 1800 2 091948
22 159 4.7 2332 5 P91365
23 159 4.6 634 4 098511
24 158 4.6 2771 2 094974
25 157.5 4.5 642 11 094970
26 157.5 4.5 670 2 085825
27 157 4.4 713 2 092826
28 156.5 4.4 1040 2 094150
29 156 4.4 481 6 097641
30 148.5 4.3 967 3 008294
31 147 4.3 1749 2 E73032
32 146 4.3 1305 5 099165
33 145.5 4.2 1275 5 076592
34 145.5 4.2 1875 5 093691
35 144.5 4.2 1001 3 005164
36 144.5 4.2 4199 2 P74440
37 142.5 4.2 641 4 094780
38 142.5 4.2 642 4 094780
39 142.5 4.2 848 2 094780
40 142.5 4.2 1589 6 062672
41 142 4.1 1180 2 094784
42 141.5 4.1 2402 2 094787
43 139.5 4.1 713 2 094787
44 139.5 4.1 786 5 021027
45 139.5 4.1 1049 2 094787

ALIGNMENTS

RESULT 1
ID 053679
AC 053679
DT 01-NOV-1996 (Trembl, 01, Created)
DI 01-NOV-1996 (Trembl, 01, Last sequence update)
DT 01-JUN-2001 (Trembl, 17, Last annotation update)
DE SIMILARITY TO CALACTINSE, OXIDASE FROM PACTYIUM PERIPOIDES.
GN FBFB GENE.
OS Stigmatella aurantiaca.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxobacteria; Cylobacteriales; Cylobacteraceae; Stigmatella.
OX NCBI.TaxID:41;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN:DM4/3-1;
FX MEDLINE 97113426, PubMed-8955286.
RA Slikowski R, Pospisil A, Neumann P, Schaller H N;
RT Stigmatella aurantiaca fruiting body formation is dependent on the
RT fefa gene encoding a polyketide homologous to chitin synthases.
RI J. Bacteriol. 178:6706-6713(1996).
DR InterPro: IPR001798; Kelch.
DR Flam. PR01344; Kelch; 5.
SQ SEQUENCE 526 AA: 57820 MW: 83356AB61DANEDG DEDG4;
GGLY Match 12.0%, Score 319.5; DB 2; Length 526;
Best Local Similarity 27.8%, Pctd. No. 3, 1e 27;
Matches 133; Conservative 79; Mismatches 139; Indels 83; Gaps 23;
148 ASSTAVAP--CICPWTITLPLVRAAAIEPTSCVLMSSSPN-AFGCTGQITIT 205
59 SSGLAAAPSEWGRMSPLMSWP-SATIAHLHSGVMF-----PDEFTGQISPR 109
206 SKEGSEIVSEFTVTKLIDMPEQISML-NDGLVVDGR--LEAFSLKSYSSDSW 260
110 LMDPLANTLT--PIRAPENIFCAGHSFTENGRLITGGVHDSIVGVDAITPKKSAW 167
261 TFCFAGVAVYSSALMSLGVVH132ASWG-CVPEKPNFVSPSTWTSTFNAAVNP 319
164 FNVFENMDEPMYNNITIANQIVV1SGETTGQGLFNEI PGPVAAVANSWNI----- 220

[illegible][illegible]

[illegible]

01 MAY 2000 (TRFBLREL_14, last annotation update)
 DE HYDROPHOBIC 61.0 KDA PROTEIN.
 GN F166190.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Euphorbiales; Euphorbiaceae; Euphorbia; Euphorbia corollata; Kosciuszko
 CC eucosmos 11; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_taxonomy:4702;
 RN 111
 RN 112
 FA Sequence from N.A.
 FA Species V. Jochmans S., Beckers J., Ausorge W., Mewes H.W.,
 FA Meyer K.E., Xie L., Lompe K., Schottorf C., Queller F., Salamounat M.,
 FA Saito T. (2004) The Arabidopsis proteome database.
 RN 121
 RN 122
 RN 123
 RN 124
 RN 125
 RN 126
 RN 127
 RN 128
 RN 129
 RN 130
 RN 131
 RN 132
 RN 133
 RN 134
 RN 135
 RN 136
 RN 137
 RN 138
 RN 139
 RN 140
 RN 141
 RN 142
 RN 143
 RN 144
 RN 145
 RN 146
 RN 147
 RN 148
 RN 149
 RN 150
 RN 151
 RN 152
 RN 153
 RN 154
 RN 155
 RN 156
 RN 157
 RN 158
 RN 159
 RN 160
 RN 161
 RN 162
 RN 163
 RN 164
 RN 165
 RN 166
 RN 167
 RN 168
 RN 169
 RN 170
 RN 171
 RN 172
 RN 173
 RN 174
 RN 175
 RN 176
 RN 177
 RN 178
 RN 179
 RN 180
 RN 181
 RN 182
 RN 183
 RN 184
 RN 185
 RN 186
 RN 187
 RN 188
 RN 189
 RN 190
 RN 191
 RN 192
 RN 193
 RN 194
 RN 195
 RN 196
 RN 197
 RN 198
 RN 199
 RN 200
 RN 201
 RN 202
 RN 203
 RN 204
 RN 205
 RN 206
 RN 207
 RN 208
 RN 209
 RN 210
 RN 211
 RN 212
 RN 213
 RN 214
 RN 215
 RN 216
 RN 217
 RN 218
 RN 219
 RN 220
 RN 221
 RN 222
 RN 223
 RN 224
 RN 225
 RN 226
 RN 227
 RN 228
 RN 229
 RN 230
 RN 231
 RN 232
 RN 233
 RN 234
 RN 235
 RN 236
 RN 237
 RN 238
 RN 239
 RN 240
 RN 241
 RN 242
 RN 243
 RN 244
 RN 245
 RN 246
 RN 247
 RN 248
 RN 249
 RN 250
 RN 251
 RN 252
 RN 253
 RN 254
 RN 255
 RN 256
 RN 257
 RN 258
 RN 259
 RN 260
 RN 261
 RN 262
 RN 263
 RN 264
 RN 265
 RN 266
 RN 267
 RN 268
 RN 269
 RN 270
 RN 271
 RN 272
 RN 273
 RN 274
 RN 275
 RN 276
 RN 277
 RN 278
 RN 279
 RN 280
 RN 281
 RN 282
 RN 283
 RN 284
 RN 285
 RN 286
 RN 287
 RN 288
 RN 289
 RN 290
 RN 291
 RN 292
 RN 293
 RN 294
 RN 295
 RN 296
 RN 297
 RN 298
 RN 299
 RN 300
 RN 301
 RN 302
 RN 303
 RN 304
 RN 305
 RN 306
 RN 307
 RN 308
 RN 309
 RN 310
 RN 311
 RN 312
 RN 313
 RN 314
 RN 315
 RN 316
 RN 317
 RN 318
 RN 319
 RN 320
 RN 321
 RN 322
 RN 323
 RN 324
 RN 325
 RN 326
 RN 327
 RN 328
 RN 329
 RN 330
 RN 331
 RN 332
 RN 333
 RN 334
 RN 335
 RN 336
 RN 337
 RN 338
 RN 339
 RN 340
 RN 341
 RN 342
 RN 343
 RN 344
 RN 345
 RN 346
 RN 347
 RN 348
 RN 349
 RN 350
 RN 351
 RN 352
 RN 353
 RN 354
 RN 355
 RN 356
 RN 357
 RN 358
 RN 359
 RN 360
 RN 361
 RN 362
 RN 363
 RN 364
 RN 365
 RN 366
 RN 367
 RN 368
 RN 369
 RN 370
 RN 371
 RN 372
 RN 373
 RN 374
 RN 375
 RN 376
 RN 377
 RN 378
 RN 379
 RN 380
 RN 381
 RN 382
 RN 383
 RN 384
 RN 385
 RN 386
 RN 387
 RN 388
 RN 389
 RN 390
 RN 391
 RN 392
 RN 393
 RN 394
 RN 395
 RN 396
 RN 397
 RN 398
 RN 399
 RN 400
 RN 401
 RN 402
 RN 403
 RN 404
 RN 405
 RN 406
 RN 407
 RN 408
 RN 409
 RN 410
 RN 411
 RN 412
 RN 413
 RN 414
 RN 415
 RN 416
 RN 417
 RN 418
 RN 419
 RN 420
 RN 421
 RN 422
 RN 423
 RN 424
 RN 425
 RN 426
 RN 427
 RN 428
 RN 429
 RN 430
 RN 431
 RN 432
 RN 433
 RN 434
 RN 435
 RN 436
 RN 437
 RN 438
 RN 439
 RN 440
 RN 441
 RN 442
 RN 443
 RN 444
 RN 445
 RN 446
 RN 447
 RN 448
 RN 449
 RN 450
 RN 451
 RN 452
 RN 453
 RN 454
 RN 455
 RN 456
 RN 457
 RN 458
 RN 459
 RN 460
 RN 461
 RN 462
 RN 463
 RN 464
 RN 465
 RN 466
 RN 467
 RN 468
 RN 469
 RN 470
 RN 471
 RN 472
 RN 473
 RN 474
 RN 475
 RN 476
 RN 477
 RN 478
 RN 479
 RN 480
 RN 481
 RN 482
 RN 483
 RN 484
 RN 485
 RN 486
 RN 487
 RN 488
 RN 489
 RN 490
 RN 491
 RN 492
 RN 493
 RN 494
 RN 495
 RN 496
 RN 497
 RN 498
 RN 499
 RN 500
 RN 501
 RN 502
 RN 503
 RN 504
 RN 505
 RN 506
 RN 507
 RN 508
 RN 509
 RN 510
 RN 511
 RN 512
 RN 513
 RN 514
 RN 515
 RN 516
 RN 517
 RN 518
 RN 519
 RN 520
 RN 521
 RN 522
 RN 523
 RN 524
 RN 525
 RN 526
 RN 527
 RN 528
 RN 529
 RN 530
 RN 531
 RN 532
 RN 533
 RN 534
 RN 535
 RN 536
 RN 537
 RN 538
 RN 53

[illegible]

```

1 APPLICATION NUMBER: US 38/469,795
2 FILING DATE: 24-MAR-1995
3 ALIBREY/AGENT INFORMATION:
4 NAME: SILVA, Robin M.
5 REGISTRATION NUMBER: 38,304
6 REFERENCE/DOC#ET NUMBER: A-61053-2/PPT/RMS
7 TELECOMMUNICATION INFORMATION:
8 TELEPHONE: (415) 781-1989
9 TELEFAX: (415) 398-3249
10 TELEX: 910 277299
11 INFORMATION FOR SEQ ID NO. 4:
12 SEQUENCE CHARACTERISTICS:
13 LENGTH: 1512 amino acids
14 TYPE: amino acid
15 STRANDNESS: unknown
16 TOPOLOGY: unknown
17 MOLECULE TYPE: protein
18 US-08-685-467-4
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
100
```

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

1794 EPIDEMIOLOGICAL SURVEILLANCE 1835

120 PARVIVATTEANOPWISIAFNWASSY-----TAPQPGIOWGPTIDPIV 170

299 -----VITPSNIP- VIKQIGAPSPYQVSNLTATQGGKVLK- DAVSA 145

171 PAAA-----LEPIS-----GRVIMSSYNDAPVNSV- 139

346 FFAVASPEVFPIDGLFVITLPPANDVAPAKVLA INWIMPTVOFLDADHNT 402

200 -CITLSSWDPSTGIVS-----DRVTYKHMF-CDJISDNGOIVT 242

403 KSVIATGPTGFTGASNNXSANFYLIPANAPVTT-----ONLVTGGGVLP 455

243 GGNVAKTSLYD-----SSSDSNIP-PMVADVGYSSATMSW- VETIAGNS 291

456 GC-----YDYCINPFAVSKMIFAGMDNAPAR-YDULFFAKRYTFMRBGM 506

232 GGVKNGEYVSSKRWGLFNKVNHLA-----KGGIYFVFNAMLFQWKGQSV 347

507 GQTDMEVEDSPASYTYVYRGTKRKGLATFEEDV- AAGNHYCVFK- 559

348 AGNSTAMWYTSSE- FVSAKRGKNGVAPDAMNNAV-----YDVAKXK 396

560 -----YTCVSPKVEEVVLS- SRFVYENLISAVGKVLKMDAPNCT- 605

397 LTFGSPDYODSDATNAHITLGE-----PITSPNV-FAS 432

606 ---PNNPNPNPT- TLESFENDIPASKTLDAIDGCHKONAPGIAQVNS 656

433 NGYFAF- HTSVLPD 449

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

657 NGVYSESGGLGVLTPD 676

PR	27 MAY-1999	9905-014382
PR	28 MAY-1999	9905-014382
PR	1 JUN-1999	9905-014382
PR	4 JUN-1999	9905-014382
PR	7 JUN-1999	9905-014382
PR	10 JUN-1999	9905-014382
PR	13 JUN-1999	9905-014382
PR	16 JUN-1999	9905-014382
PR	17 JUN-1999	9905-014382
PR	18 JUN-1999	9905-014382
PR	19 JUN-1999	9905-014382
PR	20 JUN-1999	9905-014382
PR	21 JUN-1999	9905-014382
PR	22 JUN-1999	9905-014382
PR	23 JUN-1999	9905-014382
PR	24 JUN-1999	9905-014382
PR	25 JUN-1999	9905-014382
PR	26 JUN-1999	9905-014382
PR	27 JUN-1999	9905-014382
PR	28 JUN-1999	9905-014382
PR	29 JUN-1999	9905-014382
PR	30 JUN-1999	9905-014382
PR	1 JUL-1999	9905-014382
PR	2 JUL-1999	9905-014382
PR	3 JUL-1999	9905-014382
PR	4 JUL-1999	9905-014382
PR	5 JUL-1999	9905-014382
PR	6 JUL-1999	9905-014382
PR	7 JUL-1999	9905-014382
PR	8 JUL-1999	9905-014382
PR	9 JUL-1999	9905-014382
PR	10 JUL-1999	9905-014382
PR	11 JUL-1999	9905-014382
PR	12 JUL-1999	9905-014382
PR	13 JUL-1999	9905-014382
PR	14 JUL-1999	9905-014382
PR	15 JUL-1999	9905-014382
PR	16 JUL-1999	9905-014382
PR	17 JUL-1999	9905-014382
PR	18 JUL-1999	9905-014382
PR	19 JUL-1999	9905-014382
PR	20 JUL-1999	9905-014382
PR	21 JUL-1999	9905-014382
PR	22 JUL-1999	9905-014382
PR	23 JUL-1999	9905-014382
PR	24 JUL-1999	9905-014382
PR	25 JUL-1999	9905-014382
PR	26 JUL-1999	9905-014382
PR	27 JUL-1999	9905-014382
PR	28 JUL-1999	9905-014382
PR	29 JUL-1999	9905-014382
PR	30 JUL-1999	9905-014382
PR	31 JUL-1999	9905-014382
PR	1 AUG-1999	9905-014382
PR	2 AUG-1999	9905-014382
PR	3 AUG-1999	9905-014382
PR	4 AUG-1999	9905-014382
PR	5 AUG-1999	9905-014382
PR	6 AUG-1999	9905-014382
PR	7 AUG-1999	9905-014382
PR	8 AUG-1999	9905-014382
PR	9 AUG-1999	9905-014382
PR	10 AUG-1999	9905-014382
PR	11 AUG-1999	9905-014382
PR	12 AUG-1999	9905-014382
PR	13 AUG-1999	9905-014382
PR	14 AUG-1999	9905-014382
PR	15 AUG-1999	9905-014382
PR	16 AUG-1999	9905-014382
PR	17 AUG-1999	9905-014382
PR	18 AUG-1999	9905-014382
PR	19 AUG-1999	9905-014382
PR	20 AUG-1999	9905-014382
PR	21 AUG-1999	9905-014382
PR	22 AUG-1999	9905-014382
PR	23 AUG-1999	9905-014382
PR	24 AUG-1999	9905-014382
PR	25 AUG-1999	9905-014382
PR	26 AUG-1999	9905-014382
PR	27 AUG-1999	9905-014382
PR	28 AUG-1999	9905-014382
PR	29 AUG-1999	9905-014382
PR	30 AUG-1999	9905-014382
PR	31 AUG-1999	9905-014382
PR	1 SEP-1999	9905-014382
PR	2 SEP-1999	9905-014382
PR	3 SEP-1999	9905-014382
PR	4 SEP-1999	9905-014382
PR	5 SEP-1999	9905-014382
PR	6 SEP-1999	9905-014382
PR	7 SEP-1999	9905-014382
PR	8 SEP-1999	9905-014382
PR	9 SEP-1999	9905-014382
PR	10 SEP-1999	9905-014382
PR	11 SEP-1999	9905-014382
PR	12 SEP-1999	9905-014382
PR	13 SEP-1999	9905-014382
PR	14 SEP-1999	9905-014382
PR	15 SEP-1999	9905-014382
PR	16 SEP-1999	9905-014382
PR	17 SEP-1999	9905-014382
PR	18 SEP-1999	9905-014382
PR	19 SEP-1999	9905-014382



GenPro version 4.5
Copyright (c) 1994 - 2000 Compugen Ltd.

Search time 25.91 seconds
(without alignments)
604 219 Million cell updates/sec

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

US-09-722-602-17

ALIGNMENTS

34 125.5 3.7 2429 1 YSR9_CAEEL
35 125 3.6 635 1 XYND_PAEPO
36 125 3.6 1002 1 E3P_LACIO
37 124.5 3.6 1016 1 PMPB_CHITR
38 124 3.6 677 1 Y136_MYCTU
39 123.5 3.6 1672 1 PMFB_CHLIM
40 123 3.6 940 1 NUW4_HUMAN
41 123 3.6 930 1 PM98_CHITN
42 123 3.6 1250 1 YFAL_FCOLI
43 123 3.6 1433 1 SUBE_BACSU
44 123 3.6 1754 1 PMPB_CHLIR
45 122.5 4.4 1408 1 120K_RICPT1

RESULT 1

GAOA_DACDE STANDARD: PRT: GAO AA

AC 001745:

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE GALACTOSE OXIDASE PRPCPSOP (EC 1.1.3.9) (GAO).

GN GAOA.

OS Eucalyptium dendroideum (Cladobolium dendroideum).

SC Eucalyptus, Family: Myrtaceae, Order: Myrtales, Class: Equisetum, Phylum: Equisetum, Kingdom: Equisetum.

OC Hypocotylae; Hypocotylae; Hypocotylae

OX NCBI_TaxID=5132;

BN [1]

BP SEQUENCE FROM N. A. AND PARTIAL SEQUENCE.

KX MEDLINE=92235035; PubMed=1569070;

FA McPherson M J., Ogel Z R., Stevens C E., Yadav K D S., Keen T N., Knowles P. F.;

RT "Galactose oxidase of Eucalyptium dendroideum: Gene cloning and sequence analysis."

FL 2. Biol. Chem. 267:8146-8152(1992).

BN [2]

BP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).

KX MEDLINE=9114611; PubMed=2062850;

FA McPherson M J., Ogel Z R., Stevens C E., Yadav K D S., Keen T N., Knowles P. F.;

RT "X-ray crystallographic structure of galactose oxidase from Eucalyptium dendroideum."

FL 2. Biol. Chem. 267:8146-8152(1992).

BN [3]

BP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).

KX MEDLINE=9114611; PubMed=2062850;

FA McPherson M J., Ogel Z R., Stevens C E., Yadav K D S., Keen T N., Knowles P. F.;

RT "X-ray crystallographic structure of galactose oxidase from Eucalyptium dendroideum."

FL 2. Biol. Chem. 267:8146-8152(1992).

BN [4]

ALIGNMENTS

34 125.5 3.7 2429 1 YSR9_CAEEL
35 125 3.6 635 1 XYND_PAEPO
36 125 3.6 1002 1 E3P_LACIO
37 124.5 3.6 1016 1 PMPB_CHITR
38 124 3.6 677 1 Y136_MYCTU
39 123.5 3.6 1672 1 PMFB_CHLIM
40 123 3.6 940 1 NUW4_HUMAN
41 123 3.6 930 1 PM98_CHITN
42 123 3.6 1250 1 YFAL_FCOLI
43 123 3.6 1433 1 SUBE_BACSU
44 123 3.6 1754 1 PMPB_CHLIR
45 122.5 4.4 1408 1 120K_RICPT1

RESULT 1

GAOA_DACDE STANDARD: PRT: GAO AA

AC 001745:

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE GALACTOSE OXIDASE PRPCPSOP (EC 1.1.3.9) (GAO).

GN GAOA.

OS Eucalyptium dendroideum (Cladobolium dendroideum).

SC Eucalyptus, Family: Myrtaceae, Order: Myrtales, Class: Equisetum, Phylum: Equisetum, Kingdom: Equisetum.

OC Hypocotylae; Hypocotylae; Hypocotylae

OX NCBI_TaxID=5132;

BN [1]

BP SEQUENCE FROM N. A. AND PARTIAL SEQUENCE.

KX MEDLINE=92235035; PubMed=1569070;

FA McPherson M J., Ogel Z R., Stevens C E., Yadav K D S., Keen T N., Knowles P. F.;

RT "Galactose oxidase of Eucalyptium dendroideum: Gene cloning and sequence analysis."

FL 2. Biol. Chem. 267:8146-8152(1992).

BN [2]

BP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).

KX MEDLINE=9114611; PubMed=2062850;

FA McPherson M J., Ogel Z R., Stevens C E., Yadav K D S., Keen T N., Knowles P. F.;

RT "X-ray crystallographic structure of galactose oxidase from Eucalyptium dendroideum."

FL 2. Biol. Chem. 267:8146-8152(1992).

BN [3]

BP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).

KX MEDLINE=9114611; PubMed=2062850;

FA McPherson M J., Ogel Z R., Stevens C E., Yadav K D S., Keen T N., Knowles P. F.;

RT "X-ray crystallographic structure of galactose oxidase from Eucalyptium dendroideum."

FL 2. Biol. Chem. 267:8146-8152(1992).

BN [4]

ALIGNMENTS

34 125.5 3.7 2429 1 YSR9_CAEEL
35 125 3.6 635 1 XYND_PAEPO
36 125 3.6 1002 1 E3P_LACIO
37 124.5 3.6 1016 1 PMPB_CHITR
38 124 3.6 677 1 Y136_MYCTU
39 123.5 3.6 1672 1 PMFB_CHLIM
40 123 3.6 940 1 NUW4_HUMAN
41 123 3.6 930 1 PM98_CHITN
42 123 3.6 1250 1 YFAL_FCOLI
43 123 3.6 1433 1 SUBE_BACSU
44 123 3.6 1754 1 PMPB_CHLIR
45 122.5 4.4 1408 1 120K_RICPT1

RESULT 1

GAOA_DACDE STANDARD: PRT: GAO AA

AC 001745:

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE GALACTOSE OXIDASE PRPCPSOP (EC 1.1.3.9) (GAO).

GN GAOA.

OS Eucalyptium dendroideum (Cladobolium dendroideum).

SC Eucalyptus, Family: Myrtaceae, Order: Myrtales, Class: Equisetum, Phylum: Equisetum, Kingdom: Equisetum.

OC Hypocotylae; Hypocotylae; Hypocotylae

OX NCBI_TaxID=5132;

BN [1]

BP SEQUENCE FROM N. A. AND PARTIAL SEQUENCE.

KX MEDLINE=92235035; PubMed=1569070;

FA McPherson M J., Ogel Z R., Stevens C E., Yadav K D S., Keen T N., Knowles P. F.;

RT "Galactose oxidase of Eucalyptium dendroideum: Gene cloning and sequence analysis."

FL 2. Biol. Chem. 267:8146-8152(1992).

BN [2]

BP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).

KX MEDLINE=9114611; PubMed=2062850;

FA McPherson M J., Ogel Z R., Stevens C E., Yadav K D S., Keen T N., Knowles P. F.;

RT "X-ray crystallographic structure of galactose oxidase from Eucalyptium dendroideum."

FL 2. Biol. Chem. 267:8146-8152(1992).

BN [3]

BP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).

KX MEDLINE=9114611; PubMed=2062850;

FA McPherson M J., Ogel Z R., Stevens C E., Yadav K D S., Keen T N., Knowles P. F.;

RT "X-ray crystallographic structure of galactose oxidase from Eucalyptium dendroideum."

FL 2. Biol. Chem. 267:8146-8152(1992).

BN [4]

611 - - - - - KROGSSLSLSMELHYDRIHNMWSDPMKRGKV - - - - - VALIGDHLVAVG5 657
 458 QKQKFFPENVTFEFLYPEVDTEFKQNPNSIVPAVHSHLIPDPVNRGGGLGCD 517
 658 HAPASHMDKSHLLLYEEVEFEFTIMVMVALSPSPFAVSVLLDTELVAVG3 - - - - - 710
 518 GTTHHDAVGLITFVY - - - - - YDENG 539
 711 - - - - - YDQV - - - - - TYLNMMSYDQIN 728

 EMBL 6
 GENBANK
 ID: ICEN_PSENY STANDARD; PRI: 1200 AA.
 006659;
 01-JAN-1988 (Rel. 06, last sequence update)
 01-JAN-1988 (Rel. 06, last sequence update)
 20-APR-2001 (Rel. 40, last annotation update)
 ICE_NUCLEATION PROTEIN.
 INAZ.
 Pseudomonas aeruginosa (P. aeruginosa).
 Bacterial; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 Pseudomonas
 NCBI_Acid:121;
 X
 111
 SEQUENCE FROM N.A.
 AC
 STRAIN 5203;
 AC
 Green R.L., Warren G.J.;
 KT
 "Physical and functional replication in a bacterial ice nucleation
 KT
 gene";
 1 Nature 317:645-648(1985).
 121
 SUPPLEMENTAR LOCATION.
 44
 MEDLINE 92247963; PubMed 2529825;
 44
 X
 Lindow S.E., Lindow E., Gortchakov A.G., Panopoulou N.T., Gies D.;
 44
 "Localization of ice nucleation activity and the ice gene product in
 44
 Pseudomonas syringae and Escherichia coli";
 44
 Proc. First Microbe Interact. 2:762-77(1986)
 44
 44
 3D-STRUCTURE MODELING OF 490-535.
 44
 MEDLINE 92362459; PubMed 9355257;
 44
 Kajava A.V., Lindow S.E.;
 44
 "A model of the three dimensional structure of ice nucleation
 44
 proteins";
 44
 21 J. Mol. Biol. 232:706-717(1993)
 44
 44
 1- FUNCTION: ICE NUCLEATION PROTEINS ENABLE BACTERIA TO NUCLEATE
 44
 CRYSTALLIZATION IN SUPERCOOLED WATER.
 44
 1- SUPPLEMENTAR LOCATION: OUTER MEMBRANE.
 44
 1- DOMAIN: SCYMAINS 122 IMPDET REPATE OF THE CONSENSUS
 44
 OCTAPETIDE A-C-Y-G-S-T-T-T. FUDTHER ON A 16-RESIDUE AND A
 44
 REGIONAL 48-RESIDUE DEPTIDITY IS SUPERIMPOSED.
 44
 1- MISCELLANEOUS: A STRUCTURAL MODEL IS SUGGESSED IN WHICH THE ICE
 44
 NUCLEATION PROTEIN DISPLAYS A SYMMEY RELATED TO THAT OF ICE.
 44
 1- SIMILARITY: BELONGS TO THE BACTERIAL ICE NUCLEATION PROTEIN
 44
 FAMILY.
 44

 THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 44
 BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
 44
 AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 44
 USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 44
 MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 44
 ENTERPRISES REQUIRES A LICENSE AGREEMENT (SEE <http://www.isb.sib.ch/announce/>)
 44
 OF SEND AN EMAIL TO license@sib.sib.ch.
 44

 EMBL: X59035; CNA26837.1; -;
 44
 PRI: A24405; A24405.
 44
 PRI: 11NA; 31-OCT-93.
 44
 PRI: 11NB; 31-OCT-93.
 44
 INTERPRET: PR000238, ICE_NUCLEATE.
 44
 PLAM: PR000181; ICE_NUCLEATION; 61.
 44
 FRINIS: PR00027, ICE_NUCLEATEIN.
 44
 PROSITE: PR000314; ICE_NUCLEATION; 40.
 44
 ICE_NUCLEATION: Repeat: Outer membrane; 3D-structure.
 44
 W

```

FT          TMAIN IN      176     1151    OCTAPEPTIDE PERIODICITY .
SQ          SEQUENCE     1200 AA: 118587 MW:  E3FEAFAPFA67AH9D PCFG4:

20-37 Match           4 1%   Score 142, DR 1, Length 1200;
Best local similarity 22.1%; Prod No. 9,18;
Matches 158; Conservative 87; Mismatches 341; Indels 170; Gaps 38;

CY       1 ASAPVGS---ALPENNMAVTCTGAGSGNFRKATIG--NRKTENIT     FYANCDER 50
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       158 AAFPNNILPVTTPOAVQATATGSTLSTGNHSHLLLYGYSNELTAGNSHOT IATGSGTGTAG 217
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY       51 PRHYTIIDMKTQNVNGSLVDPRQDNQNMGIAPHRYLLSDGIN---WSPVAGSMF 106
            : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       218 SDWMVAAYGSHYUIMADSSALT-----AGYSTLTAFPGSNITATYGSTTAGS-- 266
            : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY       107 AKST---TKYSNETPPARYVRIVAIITEANQGWTISIAELNFGASSTYAPPGPI/PKWGT 164
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       267 -PSSTIACVGSTQTSQRD---SSLTACVGSQTGA-----PGSNLTLAGYGSTTAGS- 314
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY       145 HTPVPAATAETPSGRGVLMWSYPN-----PAFGSEG----GIT 202
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       315 -DSLLTAHVSGHQIGDLSSSLTAGYSGTGAFGSNITAGYSGSTHAVDSSLTAGYST 373
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY       200 ITGSNWSTGTVSPPTVVTRKIMRPGTSMGNCNQYVPQQMDAKRTSYGSSNSKWP 262
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       374 QFSQSUSNI TATVSTGTAGGCSNLTAGYSTG-----TAGSESI TAGYSTQTS-- 424
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY       263 -IIIPGVAAQYSSAMSELGFHTIGSNSG-----GVFEKKGEVTFYSKTI-- 308
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       425 QLSLSLTATVSGTGQAQRTSTTATVSTGTATVNSS IATGYSTQTS---SPGATAG 481
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY       309 WTSLPNAAVNPLTA-DKQGLYPSDNIAMLEFMKKSGVDAQGPSTAMNY----- 358
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       482 YGSGTGAFGSNITATYGSTGTAGADS-SLING--GSTGTGSSSLTAGYGSTGTARE 538
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY       359 ----ISSGDVKSAAKKRS-----NQLYAPRAMQNAVMYIAVKRKITTG-GSRP 404
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       539 GSTLAGYSGSTGTAGADSSLTAGYSGSTTSGSESLTAGYSGTGTQAQGSVLISGYST- 597
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY       455 VAGSVNTINAHITTLTEPTSPNFVANSLEFEATE-ITVSULP-----DASTFI 454
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       598 -GTAAASNL-TTCYSGSTTACHESLTAGYSGTACHKSLITATAGYSGTGAARDSDLI 655
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY       465 TGQGGPIGDPEDSTVEFPPELIYEPQDTFYK-----QNPNSTVAYHSILL 501
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       656 AYVNDRIITAGSSSLIAG---YGSYLIASYPEMLTAGYSGSTFAEHLDELGYSTSTA 712
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY       592 LPHGVNPVQSLDLDELCNHHTIAGLITUNVYISNCAITPKTTPSTGVKRGARIE 561
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       713 GSNSSLIAG---YSTGTAGFKS-LITAGY-----GSTGTAGCF--PTSL-----VAGYSS 756
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY       562 ISTESSIKKASIIPVGTATHVNTQGPRTPTITNNQNSGFFGVDPSTGVALPY 617
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       757 TST-AGYSSSLIAGYG-STGTAGYES---TITAGYG---STGTAEHNSSLITGY 802
            | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT      7
ID         ALDA_ECOLI        STANDARD:      PRT: 1286 AA.
AC         Q03155.
DT         01-JUN-1994 (Rel. 29, Created)
UI         01-JUN-1994 (Rel. 29, Last sequence update)
DT         01-JUN-1984 (Rel. 29, Last annotation update)
DE         ADRESIN ALDA-I PRECUPROP.
CN         ALDA-I.
OS         Escherichia coli.
OG         Plasmid pIB6.
CC         Bacteria; Proteobacteria; gamma subdivision; Proteobacteriaceae;
CX         Escherichia
XX         NCBI_TAXID 562;
RN         [1]
RP         SEQUENCE FROM N.A., AND SEQUENCE OF 50-56.
```



```

CG          entities requires a license agreement (see http://www.istc.ch/ann-mrcc);
CC          or send an email to license@istc.ch).
-----
DR          EMBL: AF252283; AAF81719.1; -
DR          EMBL: AF252279; AAF81716.1; -
DR          EMBL: AF040943; BAA46614.1; ALT_INIT
DR          EMBL: AF353738; CAC16128.1; -
DR          MIM: 605332; -
DR          InterPro: IPR000210; BTB-POZ.
DR          InterPro: IPR001798; Kelch.
DR          Pfam: PF00651; BTB; 1.
DR          Pfam: PF01344; Kelch; 6.
DR          PRINTS: PRO6501; KETCHREPEAT.
DR          SMART: SM00335; BTB_1.
DR          PROSITE: PS50097; BTB; 1.
DR          CytoKlelelin; Actin-binding; Repeat.
KM          DOMAIN 43 88 SER-RICH.
FT          DOMAIN 212 279 BTB.
FT          REPEAT 460 506 KETCH 1.
FT          REPEAT 537 553 KILCH 2.
FT          REPEAT 555 600 KETCH 3.
FT          REPEAT 601 647 KILCH 4.
FT          REPEAT 649 700 KETCH 5.
FT          REPEAT 701 747 KETCH 6.
SQ          SEQUENCE 748 AA: 82688 MW: 113348 Da:EFPPY CRC64;

```

Matches 80; Conservative 37; Mismatches 131; Indels 74; Gaps 17;

[illegible]

RESULT	11		
SCRBL_LIMPO		STANDARD;	PRT; 916 AA.
1D	SCRBL_LIMPO		
AC	Q25386;		
ET	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV 1997 (Rel. 35, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	BETA-SCRUIN.		
OS	Eukaryotes, Polychaeta (Atlantic horseshoe crab).		
OC	Eukaryota, Metazoa, Arthropoda, Chelicerata, Merostomata; Xiphosura;		
OC	Limulidae; Limulus.		
OX	NCBI_TaxID=6850;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Testis;		
PX	METLINE=95019249; PubMed-7593276;		
VA	Way M., Sanders M., Chatelet M., Tu Y.H., Knight A., Matsudaira P.;		

"Protophylla, a homologue of the actin crosslinking protein scruin, is localized to the acrosomal vesicle of limulus sperm." J. Cell Sci. 108:3155-3162(1995).

1. FUNCTION: MAY HAVE AN ENZYMATIC ROLE. FOUND THE ACROSOMAL VESICLE AT THE ANTERIOR OF SPERM BUT NOT IN THE ACROSOMAL PROCESS.

1. TISSUE SPECIFICITY: SPERM.

1. SIMILARITY: CONTAINS 12 KELCH REPEATS.

This SWISS-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/seqdb/doc/> or send an email to license@ebi.ac.uk).

EMBL: Z17541; GABR589.1; .

InterPro: IPR001724; Kelch.

PIR: P00144; Kelch; 11.

PRINTS: PR00501; KELCHREPEAT.

Query Match 4.0% Score 138; DB 1; Length 916;
Best Local Similarity 21.8% Evid. No. 0.22;
Matches 792 Conservative 45; Mismatches 122; Indels 116; Gaps 19;
167 LPTVAALATPTTSGVIMWSSYHNAFEGSNQGLIUSMSPSGISGLIVIVIRHM 226
168 LPTVAALATPTTSGVIMWSSYHNAFEGSNQGLIUSMSPSGISGLIVIVIRHM 226
584 LPTVAALATPTTSGVIMWSSYHNAFEGSNQGLIUSMSPSGISGLIVIVIRHM 226
227 PPTQISMGN-----GQIVTIGND-----AKKISLYSSSSDWMIPGPMQ 267
228 PPTQISMGN-----GQIVTIGND-----AKKISLYSSSSDWMIPGPMQ 267
619 FETEMSLPRNHMAAYKGLIYVIGTDPHTPCWTEWATKMEFYVYSSNKWTVAADH 678
266 VAPGYSSATKSPVFTTIGSMGSGVFEKNGEYSSSKRTTSLPNAKVNMLTADKG 327
679 SASHSHSVVFNLS-IVTIGERLDSPI SASVSVVPAIDEN-----QPKPM-PIPPMG 721
428 LKSTSNHAMI FAKKFGSVFACPSIANNVYTSQSGVFSAGFPGSNQVADAMC----- 383
429 LKSTSNHAMI FAKKFGSVFACPSIANNVYTSQSGVFSAGFPGSNQVADAMC----- 383
432 MAVVSHACVY--WVGCV-----TSKQ-----NINPVLTCVCTGVV 769
694 -----GNAVYDAVKKPIIPGSSPVYQSDATTNAHITIGPGYSIN 427
770 FKHVSSKPIKIAAFSAVVD--DKTWIGGGAHNSDE--NNVLVSI--PAID-- 818
426 GYAGNMLAAT-----SLVLP-DGSEILIGGKRGIPEDSLIYVIFELYVPEQ 480
619 -VQNEALDMLOKALISCPRISSVVALESCLYILG--GINSHELSAINEELYTIDS 873

15-III-1998 (Rel. 36, Created)
BT 15-III-1998 (Rel. 36, Last sequence update)
DE 20-III-2001 (Rel. 40, Last annotation update)
DE VEGETABLE INCOMPATIBILITY PROTEIN HET-EL.
GN HET-EL.
OS Podospora anserina.
OC Eubacteria; Fori; Ascomycota; Podosporea; Cordariomycetes;
OC Sorbinales; Lasiosphaeriaceae; Podospora.
OX NCBI:taxid=5145;
RN (1)
RP SEQUENCE FROM N.A.
PA MEDLINE:9609901; PubMed-7557402;
FA Sauge S., Turgeon R., Requetet J.
RT "A gene responsible for carbohydrate incompatibility in the fungus
KT podospora anserina encodes a protein with a GTP-binding motif and G
beta homologous domain".
KL Gene 162:135-139(1995).
CC -1- FUNCTION: RESPONSIBLE FOR VESITATIVE INCOMPATIBILITY THROUGH
CC SPECIFIC INTERACTIONS WITH DIFFERENT ALLELES OF THE UNLINKED GENE,
CC HET-C.
CC -1- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).

This SWISS-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/seqdb/doc/> or send an email to license@ebi.ac.uk).

EMBL: L29125; AAB5775.1; .

InterPro: IPR001680; WD40.

PIR: P00400; WD40; 10.

PRINTS: P008320; GPOTRINBRPT.

SMART: SM00320; WD40; 10.

PROSITE: PS00678; WD_REPEATS_1; 10.

PROSITE: PS00682; WD_REPEATS_2; 10.

FT PROSITE: PS0294; WD_REPEATS_REGION; 1.

FT GTP-binding; Repeat; WD repeat.

FT NP_BIND; 300 307 GTP (POTENTIAL).

FT REPEAT; 839 869 WD 1.

FT REPEAT; 881 911 WD 2.

FT REPEAT; 923 953 WD 3.

FT REPEAT; 965 995 WD 4.

FT REPEAT; 1007 1037 WD 5.

FT REPEAT; 1049 1079 WD 6.

FT REPEAT; 1091 1121 WD 7.

FT REPEAT; 1133 1163 WD 8.

FT REPEAT; 1175 1205 WD 9.

FT REPEAT; 1217 1247 WD 10.

FT SEQUENCE 1366 AA; 149765 MW; 965483198440651 CkC64;

Query Match 4.0% Score 138; DB 1; Length 1356;
Best Local Similarity 20.5% Evid. No. 0.37;
Matches 138; Conservative 71; Mismatches 215; Indels 220; Gaps 42;
3 AATGSAIP-----NNAAVYTSQSGVFSAGFPGSNQVADAMC-----D 48
11 LKSTSNHAMI FAKKFGSVFACPSIANNVYTSQSGVFSAGFPGSNQVADAMC----- 383
802 APTDSMKYTPKKEEPWISTISVFAFWNACTQTFPHSSSVI SVAFSAIDQVVASD 861
49 KRPHTYITIMKTNNNGISVTPROENNGWIGRHYVYSSSGITMWSGVASGS----- 104
862 DK-----TRKIMPTASTGTQTLCHGGSV--W-----SVASPRER-----VASGSDUKT 906
105 ---AFADSTKYSNFTPTPAPYALVAITFANGOPMTIAF---INVGASSYAPAPGL 158
907 IKIMDAASGTCTTLEPHSGP-VLSVAFS-FEGGRVANGSTPHIKIMDAASGTCTTLE 964
159 GPWGPITDLPVWAAAALPTSEVIMWSSYHNAFEGSNQGLIUSMSPSGISGLIVIVIRHM 226
945 GHSSEVY-----SVATSPGCV-----ASGCVPTIKIMPTASTGTCTT 1002

[illegible][illegible]

KA [3]
RN SEQUENCE FROM N.A.
RP
KA Becker J.R.;
RL Submitted (OCT 1997) to the EMBL/GenBank/CCF databases
RN [4]
RN SEQUENCE FROM N.A.
RP
KA Becker J.R.;
RL Submitted (JAN 1999) to the EMBL/GenBank/CCF databases.
RN [5]
RN SEQUENCE FROM N.A.
RP
KA Becker J.R.;
RL Submitted (MAY 2000) to the EMBL/GenBank/CCF databases.
RN [6]
RN SEQUENCE FROM N.A.
RP
KA Cheuk R., Shinn P., Brooks S., Buehler E., Chao G., Johnson-Hopson C.,
KA Khan S., Kim G., Alattai H., Bel B., Chin C., Chou J., Choi E.,
KA Doni I., Conway A., Gonzalez A., Hargrove M., Haxel R., K...
KA Lee T., Lees C., Li J., Liu A., Liu J., Liu S., Mukharly N.,
KA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
KA Thavert A., Tordiani M., Weyberg M., Yu G., Davis P., Pederspiel N.,
KA Theologis A., Becker J.;
RL Submitted (AUG-2000) to the EMBL/GenBank/CCF databases
LK FIML AC002130; A000452.1;
RN SEQUENCE 615 AA; 6775 PW; 6AA51F686659FAD C0054;

DT 01-MAR-2001 (TFEMH:rel. 16, Last annotation update)
DE F10A5.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN
RP
RP SEQUENCE FROM N.A.
RA Cheuk R., Shim J., Brooks S., Buchler E., Chao Q., Johnson-Hopson C.
RA Kian S., Kim G., Altairi H., Bei B., Chin C., Chou J., Choi E.,
FA Chen L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.
RA Lee J., Leuz C., Li J., Liu A., Liu J., Liu S., Mukharly N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thavert A., Toriumi M., Vaysberg M., Yu G., Davis E., Federspiel N.,
RA Theologis A., Becker J.;
EL Submitted (JUN 2000) to the EMBL/GenBank/DDBJ databases.
LN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shim J., Brooks S., Buchler E., Chao Q., Johnson-Hopson C.
RA Kian S., Kim G., Altairi H., Bei B., Chin C., Chou J., Choi E.,
FA Chen L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.
RA Lee J., Leuz C., Li J., Liu A., Liu J., Liu S., Mukharly N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thavert A., Toriumi M., Vaysberg M., Yu G., Davis E., Federspiel N.,
RA Theologis A., Becker J.

[illegible]

```

RI Submitted (Jul-2000) to the EMBL/GenBank/DBJ databases.
DL EMBL Accession: AF047115.1
SL Accession: 547 AA, 60780 MW, 1006240c294e601d Dec04,
Query Match
      Query local similarity 8.2%, Score: 261.5, FR: 10; Length 547;
Matches: 111; Conservative: 81; Mismatches: 217; Indels: 87; Gaps:
QY 182 KVLWMS.....SYRDAFEQSPD-GITLS---SMDPSIGVSDRTVTKHEMF 227
   |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..
Db 58 RVLMTPTNGSPNLSLFPNCAKSSPDVAKSDCTAHSEYDVALNRIPLIV-OSNTW 116
QY 228 CPGISMGNQIVVGTG--NDPAKTLVSSDS---WTPGRD-KQVARGYGSAAIKSDC 281
   |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..
Db 117 GSSGAVTPDSTLQVTSGLDLEKRVKPLMCPPTGNSDTHVEVNGLAAPVYATNHILPDG 176
QY 282 RVETIGSGMSGVFEKNKEVYSPSSKTWTSLPAKYNPMLT-----ADKQGLY----RS 331
   |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..
Db 177 RQIITGGP---GGF--NYEFPK-----INAPFYSLPLSEINDPDQENMLTFVPLNT 226
QY 332 DRIHAKLEGMRKKGVSFAGGSTAMNNYTSSSGGVKSAGKQSKRGVAAPMGANAYWTA 391
   |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..
Db 227 DGMFLFANNRATILADYSTNTVTRPELPGGPPSY-----PSTISAVLL-P 273
QY 392 VKG---KILTFEGSGPYQSDATNNHILITGPPG---TSPTVFAASNLVYAFPHHS 444
   |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..
Db 274 IKMLVEVLVCGAPRGSGNI SWPRTFVKALDTCAININDVNPQVIEKMPARVWGDM 333
QY 445 VTFPQGST-FITGGGRRGIDPE-ESTVFTPEIYPEQ---DIEFKQNFNSIVRAVHSIS 499
   |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..
Db 334 MLPLDGNVLLINGSSSGTAAEIGREPVLAHDLHPDKPVGSRFEVGNPSTIPMYHSLA 393
QY 500 LILPPGAVENGCG3-----GLGSDCTINNEQDQIFPNPILYDSCNGLAIRKTKTRSTOS 553
   |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..
Db 394 TLLRDRIITVGSNSPHAFVNTGVLPTELRLFAEFSPLDYTKRYSL-RPSIVDPDPT 451
QY 554 VKVGGP-----TITSDSSISKASLIRGATHTVYNDQRIPL-TLNNGNGSYSPQ 605
   |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..
Db 452 TVNYGPIVLPPIVSGPVKSPVKVIMLPSPFTTHRSNMHQLPLVDHVSFKIGISKITE 511
QY 606 V-----PDSGVALPGYIMLFPVNSAGVSPVASTIRY 637
   |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..
Db 512 VVKRILPSSAILAPGTYVMEVYVNO-DLPESLGVKL 546

```




CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA
APPLICATION NUMBER: 09/08/409,995
FILING DATE: 24-MAR-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 18,304
REFERENCE: 409,995, A-115, LEFT
TELEPHONE: (415) 761-1989
TELEFAX: (415) 398-3249
TELEX: 910 277295
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1912 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: unknown
US-08-685-467-4

Query Match 3.8% Score 130.5 DB 1 Length 1912
Best Local Similarity 19.3% Prod No 0.084
Matches 148: Conservative 409 Mismatches 271: Indels 267: Gaps 35:

12 NMWAVTCSAGGN-----ECNKAIDSKDIFWMT-----FYANG- 47
1139 NTSVKNKGISAGNKEITNVKSLATYKDIQNTADFTQKKEHAANKANFEVEFGKCA 1258
1199 TVSAKTDNRGRHIVITLVAEAAVSDGLEKIDGKIKLKVDNTDGN-----LLYD 1309
48 -----DPRPHVYTTQKTLNVNNS-----VTRQLSQMNIHREYLSG 92
1139 NTSVKNKGISAGNKEITNVKSLATYKDIQNTADFTQKKEHAANKANFEVEFGKCA 1258
1259 TVSAKTDNRGRHIVITLVAEAAVSDGLEKIDGKIKLKVDNTDGN-----LLYD 1309
93 GINWGSVVASGSWFA--DSTTKYSNFEIRPARYKLVAVITLVAEAAVSDGLEKIDGKIKLKVDNTDGN-----LLYD 1309
1310 AIK-GASVAKGHENAVILVITLVAEAAVSDGLEKIDGKIKLKVDNTDGN-----LLYD 1309
150 SYTAPQPNQFWSEITLLEIVLAAALHETLSHVLNNSVYFVAFSPSTITLSSWP 269
1340 -----GRVY-----ROS MGLATEIDK 1358
210 STGVSDKIVIVKRNFGTQKSGHGVY-----TGNMARKTS----- 251
1059 EYAVIVFAKALNANF--VYVHLSATLGGHLEPANI ALFAVYLLTFAKSLK 1115
252 -LYNVEVSWTICDGVANVGGAGNMSGVEFLGNSNSVVEFNVEYSSTW- 309
1416 VKPDKNITFALANDLSV-----KSTVSEVSTGTNRPVVITGTEGLNARDSEKID 1470
1139 NTSVKNKGISAGNKEITNVKSLATYKDIQNTADFTQKKEHAANKANFEVEFGKCA 1258
1471 DANILNDIASLIDILNEMAFHNTNLTHTF-----FVAAVSVFVLAAQNVFC 1524
250 -ESTAMNY-----LISLNGVSAHKGQNPVAPVAMG--NAVVEAV 392
1059 EYAVIVFAKALNANF--VYVHLSATLGGHLEPANI ALFAVYLLTFAKSLK 1115
447 EPNST-----ETDQNSVTRFESVETFEITVFEDQTEFNQNSIVKAVHSII 501
1634 NKLEET-----QKEKLAHN-----SYVTLIDKDKLNDLVIAVAVANKAGMEV 1632
1634 EPNST-----ETDQNSVTRFESVETFEITVFEDQTEFNQNSIVKAVHSII 501
1634 EPNST-----ETDQNSVTRFESVETFEITVFEDQTEFNQNSIVKAVHSII 501

Query Match 3.8% Score 130.5 DB 3 Length 1912
Best Local Similarity 19.3% Prod No 0.084
Matches 148: Conservative 409 Mismatches 271: Indels 267: Gaps 35:

12 NMWAVTCSAGGN-----ECNKAIDSKDIFWMT-----FYANG- 47
1139 NTSVKNKGISAGNKEITNVKSLATYKDIQNTADFTQKKEHAANKANFEVEFGKCA 1258
1199 TVSAKTDNRGRHIVITLVAEAAVSDGLEKIDGKIKLKVDNTDGN-----LLYD 1309
48 -----DPRPHVYTTQKTLNVNNS-----VTRQLSQMNIHREYLSG 92
1139 NTSVKNKGISAGNKEITNVKSLATYKDIQNTADFTQKKEHAANKANFEVEFGKCA 1258
1259 TVSAKTDNRGRHIVITLVAEAAVSDGLEKIDGKIKLKVDNTDGN-----LLYD 1309
93 GINWGSVVASGSWFA--DSTTKYSNFEIRPARYKLVAVITLVAEAAVSDGLEKIDGKIKLKVDNTDGN-----LLYD 1309
1310 AIK-GASVAKGHENAVILVITLVAEAAVSDGLEKIDGKIKLKVDNTDGN-----LLYD 1309
150 SYTAPQPNQFWSEITLLEIVLAAALHETLSHVLNNSVYFVAFSPSTITLSSWP 269
1340 -----GRVY-----ROS MGLATEIDK 1358
210 STGVSDKIVIVKRNFGTQKSGHGVY-----TGNMARKTS----- 251
1059 EYAVIVFAKALNANF--VYVHLSATLGGHLEPANI ALFAVYLLTFAKSLK 1115
252 -LYNVEVSWTICDGVANVGGAGNMSGVEFLGNSNSVVEFNVEYSSTW- 309
1416 VKPDKNITFALANDLSV-----KSTVSEVSTGTNRPVVITGTEGLNARDSEKID 1470
1139 NTSVKNKGISAGNKEITNVKSLATYKDIQNTADFTQKKEHAANKANFEVEFGKCA 1258
1471 DANILNDIASLIDILNEMAFHNTNLTHTF-----FVAAVSVFVLAAQNVFC 1524
250 -ESTAMNY-----LISLNGVSAHKGQNPVAPVAMG--NAVVEAV 392
1059 EYAVIVFAKALNANF--VYVHLSATLGGHLEPANI ALFAVYLLTFAKSLK 1115
447 EPNST-----ETDQNSVTRFESVETFEITVFEDQTEFNQNSIVKAVHSII 501
1634 NKLEET-----QKEKLAHN-----SYVTLIDKDKLNDLVIAVAVANKAGMEV 1632
1634 EPNST-----ETDQNSVTRFESVETFEITVFEDQTEFNQNSIVKAVHSII 501
1634 EPNST-----ETDQNSVTRFESVETFEITVFEDQTEFNQNSIVKAVHSII 501


```

17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

DB 524 SCIONEZAEERHWOKEPLDSIFFEVNEHVHGEFFDM--ALLGKDMSPAATOLHAT- 680
CY 565 TIOBRITPLITNNGNVSFQVPSDSGVALPG 616
DB 661 -----DPGIALYG 688

```

```

Search completed January 3, 2002, 09:17:00
Job time: 182 sec

```

• •

17 OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 25802

Protein identification: signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

Arabidopsis thaliana

EP134405-A2

06 SEP-2000

25 FEB-2000 2000EP-0301439.
25 FEB-1999 990S-0121825.
25 MAR-1999 990S-012180.
25 MAR-1999 990S-0124548.
25 MAR-1999 990S-0125788.
25 MAR-1999 990S-0126264.
25 MAR-1999 990S-0126785.
01 APR-1999 990S-0127462.
01 APR-1999 990S-0128234.
01 APR-1999 990S-0128714.
01 APR-1999 990S-0129845.
01 APR-1999 990S-0130077.
01 APR-1999 990S-0130449.
01 APR-1999 990S-0130510.
01 APR-1999 990S-0130891.
01 APR-1999 990S-0131449.
01 APR-1999 990S-0132048.
01 APR-1999 990S-0132407.
01 APR-1999 990S-0132484.
01 APR-1999 990S-0132485.
01 APR-1999 990S-0132486.
01 APR-1999 990S-0132487.
01 APR-1999 990S-0132863.
01 APR-1999 990S-0134218.
01 APR-1999 990S-0134218.
01 APR-1999 990S-0134219.
01 APR-1999 990S-0134221.
01 APR-1999 990S-0134370.
01 APR-1999 990S-0134768.
01 APR-1999 990S-0134941.
01 APR-1999 990S-0135124.
01 APR-1999 990S-0135363.
01 APR-1999 990S-0135629.
01 APR-1999 990S-0136021.
01 APR-1999 990S-0136362.
01 APR-1999 990S-0136782.
01 APR-1999 990S-0137222.
01 APR-1999 990S-0137528.
01 APR-1999 990S-0137502.
01 APR-1999 990S-0137724.
01 APR-1999 990S-0138094.
01 APR-1999 990S-0138540.
01 APR-1999 990S-0138847.
01 APR-1999 990S-0139119.
01 APR-1999 990S-0139452.
01 APR-1999 990S-0139453.
01 APR-1999 990S-0139452.
01 APR-1999 990S-0139454.
01 APR-1999 990S-0139455.
01 APR-1999 990S-0139456.
01 APR-1999 990S-0139457.
01 APR-1999 990S-0139458.
01 APR-1999 990S-0139459.
01 APR-1999 990S-0139460.
01 APR-1999 990S-0139461.
01 APR-1999 990S-0139462.
01 APR-1999 990S-0139463.
18 JUN-1999 990S-0139750.
18 JUN-1999 990S-0139763.
18 JUN-1999 990S-0139817.
18 JUN-1999 990S-0139869.
18 JUN-1999 990S-0140353.
18 JUN-1999 990S-0140354.
18 JUN-1999 990S-0140695.
18 JUN-1999 990S-0140823.
18 JUN-1999 990S-0140991.
18 JUN-1999 990S-0141287.
18 JUN-1999 990S-0141842.
18 JUN-1999 990S-0142154.
18 JUN-1999 990S-0142055.
18 JUN-1999 990S-0142890.
18 JUN-1999 990S-0142803.
18 JUN-1999 990S-0142920.
18 JUN-1999 990S-0142927.
18 JUN-1999 990S-0143542.
18 JUN-1999 990S-0143624.
18 JUN-1999 990S-0144005.
18 JUN-1999 990S-0144085.
18 JUN-1999 990S-0144086.
18 JUN-1999 990S-0144325.
18 JUN-1999 990S-0144331.
18 JUN-1999 990S-0144332.
18 JUN-1999 990S-0144333.
18 JUN-1999 990S-0144334.
18 JUN-1999 990S-0144335.
18 JUN-1999 990S-0144352.
18 JUN-1999 990S-0144632.
18 JUN-1999 990S-0144684.
18 JUN-1999 990S-0144814.
18 JUN-1999 990S-0145086.
18 JUN-1999 990S-0145088.
18 JUN-1999 990S-0145087.
18 JUN-1999 990S-0145089.
18 JUN-1999 990S-0145192.
18 JUN-1999 990S-0145145.
18 JUN-1999 990S-0145218.
18 JUN-1999 990S-0145224.
18 JUN-1999 990S-0145276.
18 JUN-1999 990S-0145613.
18 JUN-1999 990S-0145618.
18 JUN-1999 990S-0145919.
18 JUN-1999 990S-0145951.
18 JUN-1999 990S-0146386.
18 JUN-1999 990S-0146388.
18 JUN-1999 990S-0146389.
18 JUN-1999 990S-0147038.
18 JUN-1999 990S-0147204.
18 JUN-1999 990S-0147302.
18 JUN-1999 990S-0147192.
18 JUN-1999 990S-0147260.
18 JUN-1999 990S-0147303.
18 JUN-1999 990S-0147416.
18 JUN-1999 990S-0147493.
18 JUN-1999 990S-0147935.
18 JUN-1999 990S-0148171.
18 JUN-1999 990S-0148319.
18 JUN-1999 990S-0148341.
18 JUN-1999 990S-0148565.
18 JUN-1999 990S-0148684.
18 JUN-1999 990S-0149368.
18 JUN-1999 990S-0149175.
18 JUN-1999 990S-0149902.
18 JUN-1999 990S-0149930.
18 JUN-1999 990S-0150565.
18 JUN-1999 990S-0150884.

14 MAY-1999: 990S-0144218.
PR 23-JUL-1999: 990S-0145224.
14 MAY-1999: 990S-0144219.
PR 26-JUL-1999: 990S-0145276.
14 MAY-1999: 990S-0144221.
PR 27-JUL-1999: 990S-0145913.
14 MAY-1999: 990S-0144270.
PR 27-JUL-1999: 990S-0145918.
14 MAY-1999: 990S-0144768.
PR 27-JUL-1999: 990S-0145919.
14 MAY-1999: 990S-0144941.
PR 28-JUL-1999: 990S-0145951.
20 MAY-1999: 990S-0135124.
PR 02-AUG-1999: 990S-0146386.
21 MAY-1999: 990S-0135353.
PR 02-AUG-1999: 990S-0146388.
24 MAY-1999: 990S-0135629.
PR 03-SEP-1999: 990S-0147038.
26 MAY-1999: 990S-0136021.
PR 04-AUG-1999: 990S-0147204.
27 MAY-1999: 990S-0136492.
PR 05-AUG-1999: 990S-0147302.
28 MAY-1999: 990S-0136782.
PR 05-AUG-1999: 990S-0147192.
01-JUN-1999: 990S-0137252.
PR 06-AUG-1999: 990S-0147260.
04-JUN-1999: 990S-0137528.
PR 06-AUG-1999: 990S-0147303.
07-JUN-1999: 990S-0137724.
PR 09-AUG-1999: 990S-0147416.
08-JUN-1999: 990S-0138094.
PR 09-AUG-1999: 990S-0147935.
10-JUN-1999: 990S-0138540.
PR 10-AUG-1999: 990S-0148171.
14-JUN-1999: 990S-0139119.
PR 11-AUG-1999: 990S-0148419.
15-JUN-1999: 990S-0139452.
PR 12-AUG-1999: 990S-0148441.
16-JUN-1999: 990S-0139453.
PR 13-AUG-1999: 990S-0148565.
17-JUN-1999: 990S-0139492.
PR 14-AUG-1999: 990S-0148684.
18-JUN-1999: 990S-0139494.
PR 16-AUG-1999: 990S-0149368.
18-JUN-1999: 990S-0139495.
PR 17-AUG-1999: 990S-0149175.
18-JUN-1999: 990S-0139496.
PR 18-AUG-1999: 990S-0149426.
18-JUN-1999: 990S-0139497.
PR 20-AUG-1999: 990S-0149722.
18-JUN-1999: 990S-0139498.
PR 20-AUG-1999: 990S-0149723.
18-JUN-1999: 990S-0139499.
PR 20-AUG-1999: 990S-0149929.
18-JUN-1999: 990S-0139499.
PR 24-AUG-1999: 990S-0149902.
18-JUN-1999: 990S-0139499.
PR 24-AUG-1999: 990S-0149930.
18-JUN-1999: 990S-0139499.
PR 25-AUG-1999: 990S-0150566.
18-JUN-1999: 990S-0139499.
PR 26-AUG-1999: 990S-0150884.
18-JUN-1999: 990S-0139499.
PR 27-AUG-1999: 990S-0151065.
18-JUN-1999: 990S-0139499.
PR 27-AUG-1999: 990S-0151066.
18-JUN-1999: 990S-0139499.
PR 30-AUG-1999: 990S-0151080.
18-JUN-1999: 990S-0139499.
PR 31-AUG-1999: 990S-0151303.
18-JUN-1999: 990S-0139499.
PR 01-SEP-1999: 990S-0151930.
18-JUN-1999: 990S-0139499.
PR 07-SEP-1999: 990S-0152363.
18-JUN-1999: 990S-0139499.
PR 10-SEP-1999: 990S-0153070.
18-JUN-1999: 990S-0139499.
PR 13-SEP-1999: 990S-0153758.
18-JUN-1999: 990S-0139499.
PR 15-SEP-1999: 990S-0154018.
18-JUN-1999: 990S-0139499.
PR 16-SEP-1999: 990S-0154039.
18-JUN-1999: 990S-0139499.
PR 20-SEP-1999: 990S-0154779.
18-JUN-1999: 990S-0139499.
PR 22-SEP-1999: 990S-0155189.
18-JUN-1999: 990S-0139499.
PR 24-SEP-1999: 990S-0155486.
18-JUN-1999: 990S-0139499.
PR 28-SEP-1999: 990S-0155659.
18-JUN-1999: 990S-0139499.
PR 29-SEP-1999: 990S-0156458.
18-JUN-1999: 990S-0139499.
PR 04-OCT-1999: 990S-0157117.
18-JUN-1999: 990S-0139499.
PR 05-OCT-1999: 990S-0157853.
18-JUN-1999: 990S-0139499.
PR 06-OCT-1999: 990S-0157865.
18-JUN-1999: 990S-0139499.
PR 07-OCT-1999: 990S-0158202.
18-JUN-1999: 990S-0139499.
PR 08-OCT-1999: 990S-0158219.
18-JUN-1999: 990S-0139499.
PR 12-OCT-1999: 990S-0158469.
18-JUN-1999: 990S-0139499.
PR 13-OCT-1999: 990S-0159293.
18-JUN-1999: 990S-0139499.
PR 14-OCT-1999: 990S-0159294.
18-JUN-1999: 990S-0139499.
PR 14-OCT-1999: 990S-0159295.
18-JUN-1999: 990S-0139499.
PR 14-OCT-1999: 990S-0159329.
18-JUN-1999: 990S-0139499.
PR 14-OCT-1999: 990S-0159330.
18-JUN-1999: 990S-0139499.
PR 14-OCT-1999: 990S-0159331.
18-JUN-1999: 990S-0139499.
PR 14-OCT-1999: 990S-0159647.
18-JUN-1999: 990S-0139499.
PR 14-OCT-1999: 990S-0159648.
18-JUN-1999: 990S-0139499.
PR 18-OCT-1999: 990S-0159584.
18-JUN-1999: 990S-0139499.
PR 21-OCT-1999: 990S-0160741.
18-JUN-1999: 990S-0139499.
PR 21-OCT-1999: 990S-0160767.
18-JUN-1999: 990S-0139499.
PR 21-OCT-1999: 990S-0160770.
18-JUN-1999: 990S-0139499.
PR 21-OCT-1999: 990S-0160778.
18-JUN-1999: 990S-0139499.
PR 21-OCT-1999: 990S-0160814.
18-JUN-1999: 990S-0139499.
PR 21-OCT-1999: 990S-0160815.
18-JUN-1999: 990S-0139499.
PR 22-OCT-1999: 990S-0160980.
18-JUN-1999: 990S-0139499.
PR 22-OCT-1999: 990S-0160981.


```

10-SEP-1999 9905-0153070.
PR 14-SEP-1999 9905-0153758.
PR 15-SEP-1999 9905-0154018.
PR 16-SEP-1999 9905-0154039.
PR 20-SEP-1999 9905-0154779.
PR 22-SEP-1999 9905-0155139.
PR 23-SEP-1999 9905-0155486.
PR 24-SEP-1999 9905-0155659.
PR 28-SEP-1999 9905-0156458.
PR 29-SEP-1999 9905-0156596.
PR 04-OCT-1999 9905-0157117.
PR 05-OCT-1999 9905-0157753.
PR 06-OCT-1999 9905-0157865.
PR 07-OCT-1999 9905-0158029.
PR 08-OCT-1999 9905-0158232.
PR 12-OCT-1999 9905-0158369.
PR 13-OCT-1999 9905-0159293.
PR 14-OCT-1999 9905-0159294.
PR 14-OCT-1999 9905-0159495.
PR 14-OCT-1999 9905-0159929.
PR 14-OCT-1999 9905-0159930.
PR 14-OCT-1999 9905-0159931.
PR 14-OCT-1999 9905-0159937.
PR 14-OCT-1999 9905-0159938.
PR 18-OCT-1999 9905-0159584.
PR 21-OCT-1999 9905-0160741.
PR 21-OCT-1999 9905-0160767.
PR 21-OCT-1999 9905-0160768.
PR 21-OCT-1999 9905-0160770.
PR 21-OCT-1999 9905-0160814.
PR 21-OCT-1999 9905-0160815.
PR 22-OCT-1999 9905-0160980.
PR 22-OCT-1999 9905-0160981.
PR 22-OCT-1999 9905-0160989.
PR 25-OCT-1999 9905-0161404.
PR 25-OCT-1999 9905-0161405.
PR 25-OCT-1999 9905-0161406.
PR 26-OCT-1999 9905-0161359.
PR 26-OCT-1999 9905-0161360.
PR 26-OCT-1999 9905-0161361.
PR 28-OCT-1999 9905-0161920.
PR 28-OCT-1999 9905-0161992.
PR 28-OCT-1999 9905-0161993.
PR 29-OCT-1999 9905-0162142.

Query Match 8.387 Score 284.57 DB 21; Length 504;
Best Local Similarity 25.28; Pred. No. 6,46-13;
Matched 124; Constructive 62; Mismatched 203; Delete 107; Gaps 23;

10-SEP-1999 9905-0140823.
PR 14-SEP-1999 9905-0140991.
PR 15-SEP-1999 9905-0141287.
PR 16-SEP-1999 9905-0141842.
PR 20-SEP-1999 9905-0142055.
PR 22-SEP-1999 9905-0142390.
PR 24-SEP-1999 9905-0142803.
PR 28-SEP-1999 9905-0142920.
PR 29-SEP-1999 9905-0142977.
PR 04-OCT-1999 9905-0143542.
PR 05-OCT-1999 9905-0143624.
PR 06-OCT-1999 9905-0144005.
PR 07-OCT-1999 9905-0144085.
PR 08-OCT-1999 9905-0144306.
PR 12-OCT-1999 9905-0144325.
PR 13-OCT-1999 9905-0144331.
PR 14-OCT-1999 9905-0144332.
PR 14-OCT-1999 9905-0144333.
PR 14-OCT-1999 9905-0144334.
PR 14-OCT-1999 9905-0144375.
PR 20-SEP-1999 9905-0144352.
PR 20-SEP-1999 9905-0144632.
PR 21-SEP-1999 9905-0144864.
PR 21-SEP-1999 9905-0144884.
PR 21-SEP-1999 9905-0145086.
PR 21-SEP-1999 9905-0145088.
PR 22-SEP-1999 9905-0145085.
PR 22-SEP-1999 9905-0145087.
PR 22-SEP-1999 9905-0145089.
PR 22-SEP-1999 9905-0145192.
PR 23-SEP-1999 9905-0145145.
PR 23-SEP-1999 9905-0145218.
PR 23-SEP-1999 9905-0145221.
PR 23-SEP-1999 9905-0145276.
PR 27-SEP-1999 9905-0145918.
PR 27-SEP-1999 9905-0145919.
PR 28-SEP-1999 9905-0145951.
PR 02-AUG-1999 9905-0146386.
PR 02-AUG-1999 9905-0146388.
PR 02-AUG-1999 9905-0146389.
PR 04-AUG-1999 9905-0147038.
PR 04-AUG-1999 9905-0147204.
PR 04-AUG-1999 9905-0147302.
PR 05-AUG-1999 9905-0147192.
PR 05-AUG-1999 9905-0147260.
PR 06-AUG-1999 9905-0147403.
PR 06-AUG-1999 9905-0147416.
PR 09-AUG-1999 9905-0147493.
PR 09-AUG-1999 9905-0147935.
PR 10-AUG-1999 9905-0148171.
PR 11-AUG-1999 9905-0148319.
PR 12-AUG-1999 9905-0148341.
PR 13-AUG-1999 9905-0148565.
PR 14-AUG-1999 9905-0148684.
PR 16-AUG-1999 9905-0149368.
PR 17-AUG-1999 9905-0149375.
PR 18-AUG-1999 9905-0149426.
PR 20-AUG-1999 9905-0149722.
PR 20-AUG-1999 9905-0149723.
PR 20-AUG-1999 9905-0149724.
PR 24-AUG-1999 9905-0149902.
PR 24-AUG-1999 9905-0149930.
PR 26-AUG-1999 9905-0150566.
PR 26-AUG-1999 9905-0150884.
PR 27-AUG-1999 9905-0151065.
PR 27-AUG-1999 9905-0151066.
PR 27-AUG-1999 9905-0151080.
PR 28-AUG-1999 9905-0151353.
PR 31-AUG-1999 9905-0151438.
PR 31-SEP-1999 9905-0151620.
PR 07-SEP-1999 9905-0152363.

```


Sun Jan 6 15:49:58 2002

us-09-722-602-21.rag

Page 18

Received: April 10, 2002, 09:24:48
Accepted: May 14, 2002

us-09-722-602-21
 Copyright (c) 1993 - 2000 Computer Ltd.
 1993-09-722-602-21

Search time 45.06 seconds
 (without alignments)
 1245.536 Million cells updates/sec

Protein: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

Search: 1445
 Accession: A38084
 Description: 1445

30 140.5 4.1 1651 2 J01340
 31 140.5 4.1 2383 2 D64962
 32 140 4.1 1356 2 T18521
 33 139.5 4.1 1649 2 C86822
 34 139.5 4.1 4199 2 S76412
 35 139 4.1 786 2 T16509
 36 137.5 4.0 1217 2 T25994
 37 137 4.0 1147 2 T40866
 38 137 4.0 1286 2 S28634
 39 136.5 4.0 1529 2 T42222
 40 136.5 4.0 1662 2 T18540
 41 135.5 4.0 1300 2 S07575
 42 135.5 4.0 1461 2 A85547
 43 135 4.0 1519 2 S4125
 44 135 3.9 1797 2 T21889
 45 135 3.9 1805 2 T21888

ALIGNMENTS

RESULT 1

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

Accession: A38084
 Species: Cladobotryum dendroides
 Date: 29 Oct 2000 #sequence_version 20-Oct-2000 #text_change 20-Oct-2000

237 32AVYVGR--NAKSTSL-----YSSSSSWIPGPD--NOVAFGVSATMSDQVFTIQ 288
 126 GILVVGGRNNGFAAMSEPSDSDSD--WLETVLSNFRWATNLTPIGPIVW 314
 285 --SWSNVFPRKEVSPSSK-----TWSTLFAKVEN--LTADKGLRSDNAHLF 338
 146 PQGVNFI PPRH--DSRKRSLRLRLRLKRLSOSNRNRLNLTPIHITFGNM--VYANPESLVE 242
 339 GMRKGSVQVQANSTAMWMTTSSGSLVAFKRGQNRQVAFAMQGNVNT--DA-- 391
 244 DFFPNFVVEEFEEF-----SNNLRN-----LPSN--GSSILPPLDITNAN 283
 492 VKKLLIFKASPRGYS-----LAFTHAHTITTPPTSPVAFHPTTAPRTITVYL 117
 284 VEETIVAGGSPKPGKSPKSTBATSTGSPKLSQSPS-----WERTPIIPVMDMLL 339
 448 PQGSFTITRQRKGFEEEST--PVFTPELVPEQDTYKQNPNSIVRAVHSTILLEPDG 505
 340 PPTVITVWAGACACAKRKAATLGLVAVGRLHLLVMAITPRKMTSSAILLFDG 399
 506 PVENVGG--FGDGTTHDFACT--ETTRITGSRNMAITPRITSTQSVKAGR 559
 400 RVLVGSNHNHYNTINVEPIPLDSEAVSPPLFTSDPI--PPKILLISDKVLSY--K 455
 550 ITTSDSSISK-----ASLRYGIAHIVRDKNRPILT--TNMGNSYSKVV-- 506
 456 RUFNMTSIAGLTWLLSVETVAEN--LHSEAMNQMLKLKLSVRLDILNSTRVSA 515
 607 -PDSGVATGTMMLFVMAVGVSVASILLK 637
 516 GSTAEIAPGQIMFIVH-AGIPSSAAWVQI 546

RESULT 4
 45495
 Probable arabinose oxidase (EC 1.4.99.15) [similarity] - Arabidopsis thaliana
 (Species: Arabidopsis thaliana (mouse-ear cress))
 Protein: A1-FH-2000-1-seq-2000-1-revision 02-Mar-2001 #taxi_change 31-Mar-2001
 Accession: 145935
 Submitted to: GeneBank
 Submitted by: The Protein Sequence Database, January 2000
 Reference number: Z23017
 Accession: 145935
 Molecule type: DNA
 Residues: 1-545 a.a.
 Cross-references: EMBL:AL132960
 Experimental source: cultured Columbia, BAC clone F5K20

Query Match 7/87 Score 266, DB 2, Length 545
 Best local similarity 47/87, Pred. No. 2, to 69
 Matches 126, Conservative 67, Mismatches 141, Indels 80, Gaps 27
 225 EMPNGLQGMQNVYVIG--MAAATSLD-----SSSHKFGHBMVANG--YQSSA 276
 113 FTWSSVGRSIPSL--TQGRKPRRTAFL--LHLL--DQVLCGLLHGNKWSAW 172
 277 EMGQVAVTSSSSCH--LKNVTVGRSTW--SLIHANVNM--LAVVDLYKSNHAM 330
 173 FTGQVAVVSGDNTV--ETTRN----NCAVTVLAVVD--KQMDN-- 216
 337 FTGQVAVVSGDNTV--ETTRN----NCAVTVLAVVD--KQMDN-- 216
 217 TVVYVHITPDGSHLTPFAKRAVTHH--MAVNLH--LQDQKNTFSG--GSSAML-- 273

QY 385 NAWMYDAVKNKILTEGSSPDVUSDAITLHII--TLSP--PQISNIVFASNGLYFAKTH 442
 DB 274 -ALVGLFTLAILLGGAGSAGTARAIIDAPAHSTGRIYATADPWPVTEEMPEGRIMG 332
 QY 443 TSVLPFGSTFTFGGRRGIP--PE--DSTPVFTPELVPEOD---TFKQNPNSIVRAVS 497
 DB 333 DWNVLTPLGLILLNGAGSQQGFHM--SDPLVPLTYPPOICLRFMTLNGCYPRKHS 392
 QY 498 ISILLPDGPRVFNNGGGLGCTTNH---FDAQI-----FTPNLYNSGNLAIKPK 545
 DB 393 TANLPPDGRILLAG-----SNPHYFKFNAEPTELRIEAFSPVLSPPRANL--RPE 443
 QY 546 IIRISTQVAVGGRITITSDSSISKSLR--VGT---ATHVNTDQFRPLTL---TNN 597
 DB 444 I-GEIPQIIRYGVLEVYVILPVVGILOMNGSAPFALHSHSOGRLKVLVAVSPVD 502
 QY 598 GGNYSYFQ--VPSDGVALLPGWMLFVMAVGVSPVASTIRV 637
 DB 503 GVGRRYDQCTAPPNGAVSPGYMAFVAVNO--GVPSIARWIRI 543

RESULT 5
 H86278
 hypothetical protein AAF4943.1 [imported] - Arabidopsis thaliana
 (Species: Arabidopsis thaliana (mouse-ear cress))
 CDate: 02-Mar-2001 #sequence_revision 02-Mar-2001 #taxi_change 31-Mar-2001
 Accession: H86278
 R:Thellogis, A.; Ecker, T.P.; Platt, C.J.; Federfeld, N.A.; Kaul, S.; White, O.; Alonso, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Liu, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lopus, J.S.; Mulli, R.; Marcia, R.; Mee, M.; Penney, T.; Powley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo, K.; Wu, D.; Yu, C.; Fraser, C.M.; Venturi, C.J.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: Arabid; MIMD:21016719
 A:Accession: H86278
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-564 (SD)
 A:Cross-references: AF086172, N1-97-0655, P1DN:AAF43943.1, GSI:DB_GNC0141
 C:Genetics:
 A:Map position: 1
 C:Supplementary: Arabidopsis thaliana hypothetical protein F5K20.250

Query Match 7/18 Score 241.5, DB 2, Length 564
 Best local similarity 23/28, Pred. No. 8, to 68
 Matches 126, Conservative 77, Mismatches 224, Indels 115, Gaps 26
 QY 197 SPGGTITSSMD--PSTGT-----VSTPT-----VTV--- 221
 DB 33 SPALVTGRMDLLQPSVGSISAMHOLLHNKKVYIFDKTDGFSNVSLPSQTCQATVPD 92
 QY 222 -----TKHMECPGISMG--GNQIVVTS--GNLAKTSLY----- 253
 DB 92 CSAHSLLYDVASNTFRPLTLRYDTWSSGSI--NANST--IGTSGYSGEPLVAVFTTQDGV 152
 QY 254 PSSSTSMIPF--FHMVAVH--YSSATMS--GEVFIIGS--SWAGCVLRN--GE--VSSRSKT 368
 DB 153 GVSQGDITENPVAVLSRWSVSTNCLTPGPTITVGPVAFVHYFKPRGGEVFM--PLIA 212
 QY 309 WTSLEPAKVN--PMLTFDKQCIYPSDNIAMLFCKRKGSEVQAGSTAMNMYTSSGVK 366
 DB 213 ETDPDPENNILYFTH-----LIPDGN--ETFANPSTLTFVNHETIKETPQJLQYAKK 266
 QY 367 SAKRQSNPQVADPAMQGNVMPVAVKSKLITFGSSPYG--DSTPATIDAHITIGE--- 421
 DB 267 NYPSTGSS--VILPFLITGD--INFKRTIAEVWVGGAPGAPFAKAFITPKIFVAGSSPTG 324
 QY 422 --PQTSNIVFASNGLYFAKRAVTHSVLPDQSTFTTGGQRGIP--PEST--PVFTPELV 477

Matches 140; Conservative 84; Mismatches 275; Indels 151; Gaps 34;

```
QY 16 VTQDSASGSECKNAIDKNDTPHHFYGANGDKPKPHITYIDIKKTQNVNGLSM-LPRQ 74
Db 187 LTSSSTGCAVTPCGNGNCGNATYEWESANNA-----LPOWLOADGASARKRVRLRLP-- 239
QY 75 EGNQNGNIGRHE---VYLSSDGTNMGSFVAGSGWFADSTIKYS--NFEIRPAFVRLVA 128
Db 240 ---DGPAPRSQTLIKQASINQSDIPLTLAAKATPDAAGQGSATVFEDATATAYVR-VL 294
QY 129 ITEANQCPWTSIAEINVEQFQSSATAPQGLGRNGPTTIDIPVPAALAI-EPTSKRV-LMW 186
Db 295 VIANIQCPAPQISELEY-----GPETG-----DTQAPTPANLAFTBPATGQIRLTW 342
QY 187 SSTRNDAFGGSGPGITLTSSWDPSTGIVSDT-----VTITKDHMFCPGISMONG 237
Db 343 NESSDPT---AVTGQVYVANGDILTSVAGGVITYTIDTPRAGITYIVYHRAKDAAGNOSGD 399
QY 238 QIVTQGNPAKRTSLYUSSE-----SWTPGFLMWVAKGYUSALMSDQKV- 483
Db 400 SNSVTRAPRTGDTQATPQSLNLALEPVQGVKLWTASTIDEGCVGTG---VYADNKLK 456
QY 284 FTIGGSGSCGVFEKNGEYVS--FSKRTWTSLEPNANVPMILADKQGLYKSD-NHAWLFG 339
Db 457 KSVAG-----NVTYTDTPQASANYFEVRAK-----DAAGNEGSDNSVFRGG 500
QY 340 WRKSGVFOAG-PSIA---NMWYTSGSGCPVKSAGKQSNKNGVAPDAM---CGANWMTDAV 392
Db 501 IMDGSLNLAQKPTLTASSHITHTYVAENANQNSITWEGAGAGTPNTLTIKLGANADVDV 560
QY 393 KKKILTFGSPDYQSDATILAHILTLG---EPGTSPTNVFASNGLYPAPFTSVLPD 449
Db 561 VKL-----NPACMSRRP--QHTFVLQFAQASGLTSLVAADKYVDFDASGNSVITPV 613
QY 450 G-----STFTTCGQRGRIFFE-DSIPVFPEIYVPEQDTFYKQNPISIVRAH 456
Db 614 GKRAVDVQLKVISNIGATAGQ--VAEFQYIGTPAPRPDLEV-----TGLTSPQSPVES-- 665
QY 437 STSLIPGPRVFNQSGGLGQDCTNHPDAQITPNTLYNSNGNLATRPKITRISTQSV-- 554
Db 666 --DELTVASIVYNSGAA-----DAPA-----GKVAARLQDTKYATADVPA 703
QY 555 -KVGGRITITDSSITSKASLIRYGTATHVNTDQRIPLTLNNGNSYS 603
Db 704 LEAGQRTVSASLGAREAGSYELSAVADEAN-----ELIEQNEFNNTYT 747

RESULT 14
T34434
hypothetical protein K06A9.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 27-Oct-1999 #sequence_revision 29-Oct 1999 #text_change 18-Feb-2000
C:Accession: T34434
R:Gene(s), C: Gattung, S.
Submitted to the EMBL data library, December 1996
A:Description: The sequence of C. elegans cosmid K06A9.
A:Reference number: 231525
A:Accession: T34434
A>Status: preliminary; translated from GH/EMBL/DBJ
A:Molecule type: DNA
A:Features: 1-2342 CDS
A:Cross-reference: EMBL:NR0846, F11N AA-70650.1, GSFH:NR0808, CESP: K06A9.1a
A:Experimental source: strain Bristol N2, clone K06A9
C:Gene(s):
A:Name: CESP:K06A9.1a
A:Map position: X
A:Features: 36/37 75/79 106/66 142/23 156/23 221/1 1098/1 1667/1 2099/1 2343/1 20
```




GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

SUMMARIES

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

ALIGNMENTS

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

us-09-722-602-21.rsp

SMALL: SM00241: FAB002: 1.
Sequence: 1 41 680
CHAIN 42 680
REPEAT 279 321 GALACTOSE OXIDASE.
REPEAT 323 372 KETCH 1.
REPEAT 492 544 KETCH 2.
REPEAT 59 68 KETCH 3.
THIO-ETH 269 413
LIS-ETH 556 559
METAL 313 413 COPPER.
METAL 536 536 COPPER.
METAL 537 537 COPPER.
METAL 622 622 COPPER.
STRAND 48 49
STRAND 55 59
STRAND 63 63
TURN 64 45
STRAND 66 46
HE-IX 72 72
TURN 76 77
STRAND 80 81
HE-IX 85 87
STRAND 94 112
TURN 120 121
STRAND 123 123
STRAND 125 131
STRAND 142 145
STRAND 153 159
STRAND 172 172
TURN 173 174
STRAND 180 181
STRAND 183 188
TURN 197 198
STRAND 201 207
STRAND 214 217
TURN 219 221
STRAND 224 227
TURN 246 247
STRAND 243 249
TURN 253 253
STRAND 254 260
TURN 262 263
STRAND 271 274
TURN 276 277
STRAND 279 283
TURN 285 286
TURN 288 289
STRAND 291 295
TURN 296 299
STRAND 300 303
STRAND 312 312
TURN 315 318
STRAND 320 321
STRAND 324 327
STRAND 330 330
STRAND 332 332
TURN 340 344
TURN 345 348
STRAND 349 353
TURN 354 355
HE-IX 359 361
TURN 366 368
HE-IX 369 372
TURN 373 373
STRAND 378 380
HE-IX 382 384
STRAND 385 386
TURN 394 399
TURN 402 403
STRAND 405 411
STRAND 413 414
TURN 415 416

PT STRAND 417 418
PT STRAND 423 423
PT TURN 424 424
PT STRAND 426 431
PT TURN 432 435
PT STRAND 436 440
PT STRAND 443 443
PT STRAND 450 450
PT STRAND 452 452
PT STRAND 455 459
PT TURN 463 464
PT STRAND 468 471
PT TURN 473 474
PT STRAND 480 481
PT STRAND 484 487
PT TURN 489 490
PT STRAND 493 496
PT STRAND 499 499
PT STRAND 501 501
PT TURN 504 505
PT STRAND 510 510
PT STRAND 515 518
PT TURN 519 522
PT STRAND 523 526
PT TURN 536 537
PT STRAND 538 542
PT TURN 544 545
PT STRAND 548 552
PT TURN 557 558
PT STRAND 565 570
PT HE-IX 572 574
PT STRAND 575 575
PT TURN 577 578
PT STRAND 581 581
PT STRAND 587 590
PT STRAND 594 596
PT TURN 597 598
PT STRAND 600 605
PT STRAND 611 616
PT TURN 619 621
PT STRAND 622 623
PT STRAND 624 624
PT STRAND 630 633
PT STRAND 636 637
PT STRAND 642 646
PT TURN 651 653
PT STRAND 657 664
PT TURN 666 667
PT STRAND 670 670
PT STRAND 674 679
SQ SEQUENCE 680 AA: 72823 MW: 297056166 IE46E9 CR064:

Query Match 99.7%, Score 3416, DB 1: length 680;
Best local Similarity 99.7%; pred No 70-207;
Matches 637; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
DB 1 121 ASAPISGSAISRNMMVATCDSASGNECNKAIDGKDTWHTFYGANDPRPHYYITIMK 60
DB 12 ASAPISGSAISRNMMVATCDSASGNECNKAIDGKDTWHTFYGANDPRPHYYITIMK 101
DB 121 ASAPISGSAISRNMMVATCDSASGNECNKAIDGKDTWHTFYGANDPRPHYYITIMK 180
DB 162 ASAPISGSAISRNMMVATCDSASGNECNKAIDGKDTWHTFYGANDPRPHYYITIMK 221
DB 181 ASAPISGSAISRNMMVATCDSASGNECNKAIDGKDTWHTFYGANDPRPHYYITIMK 240
DB 222 ASAPISGSAISRNMMVATCDSASGNECNKAIDGKDTWHTFYGANDPRPHYYITIMK 281

MEHLIN-92157802: PubMed:1724278;
 RA Almore R.D. Jr., Clepik R. Jr., Policastro F.F., Hackstadt T.;
 RA "The 120 kilodalton outer membrane protein (omp B) of Rickettsia
 RA rickettsii is encoded by an unusually long open reading frame.
 RA Evidence for protein processing from a large precursor.";
 RA Mol. Microbiol. 5:3361-3370(1991).
 RN (2)
 PP REFERENCE: 279-1524 IN M.B.A.
 PP RICKETTSIAE OMPB/OMPB FAMILY.
 RX MEDLINE:90136987: PubMed:2515418;
 RX Almore R.D. Jr., Joste N., McDonald G.A.;
 RX Cloning, expression and sequence analysis of the gene encoding the
 RX 129 kD surface-exposed protein of Rickettsia rickettsii.";
 RX Mol. Microbiol. 3:1575-1586(1991).
 RX -1- FUNCTION: THE 120 KDA SURFACE-EXPOSED PROTEIN IS A MAJOR
 RX STRUCTURAL PROTEIN WHICH HAS A ROLE AS A RICKETTSIAL
 RX VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
 RX -1- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
 RX LAYER WITH HEXAGONAL SYMMETRY.
 RX -1- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPB/OMPB FAMILY.
 RX
 RX THIS SWISS-PROT ENTRY IS DELETED. It is produced through a collaboration
 RX between the Swiss Institute of Bioinformatics and the EMBL outstation
 RX the European Bioinformatics Institute. There are no restrictions on its
 RX use by non-profit institutions as long as its content is in no way
 RX modified and this statement is not removed. Usage by and for commercial
 RX entities requires a license agreement (See http://www.sdb.bio.ox.ac.uk/
 RX or send an email to license@sdb.bio.ox.ac.uk).
 RX
 RX EMBL: X16353: CAA34403.1;
 RX InterPro: IPR003656: RomA_TompB;
 RX Pfam: PF02708: rompa_tompb; 1;
 RX Antigen: S-layer; Cell wall.
 RX CHAIN 1 1333 120 KDA SURFACE-EXPOSED PROTEIN.
 RX DOMAIN 1334 1554 32 KDA BETA PEPTIDE.
 RX DOMAIN 1181 1188 POLY-1HR.
 RX SEQUENCE 1654 AA: 168184 MW: 17806870676418 98064;

Query Match 3.9%; Score 133.5; DB 1; Length 1654;

Best Local Similarity 20.1%; Pred. No. 0.89;

Matches 137; Conserving 95; Mismatches 268; Indels 183; Gaps 32;

```

74 QDGNQNGWIG RHEVLLSAGT-----NMGSP---VAGSWFADSTK 112
      | | | | | | | | | | | | | | | | | | | | | | | |
358 QSANATGOVNHIVAGAGITAKTAAKSVIIIDDSNFGNTDFGLAOLKVPNAITL 417
      | | | | | | | | | | | | | | | | | | | | | | | |
113 YSNF---EIRARVEIVAIIEANLQPTSLAEINFEQASSTAPD---PGLGKQPT-- 164
      | | | | | | | | | | | | | | | | | | | | | | | |
418 TGNFTGDSNPNINGVITF-DANITLESASADANAVATNNTAIFASGAGVVLSTHA 476
      | | | | | | | | | | | | | | | | | | | | | | | |
165 IDPIVPAALAEPIISCVLMASSTRDAFGS-----PGIITTSWDPSTGIUSD 216
      | | | | | | | | | | | | | | | | | | | | | | | |
477 AELRLQNGSIFRLADGIVINQKYNQALVAGALAGIITLDSATITGIDGNAGGAAL 536
      | | | | | | | | | | | | | | | | | | | | | | | |
217 RIVTV---IKHMFPGISMEGNS QIVTGSNDARKISLYDSSDSWIPGPMQVA 269
      | | | | | | | | | | | | | | | | | | | | | | | |
537 QRTIANAKKTIILIGCANIIGAGGTTDLQANGS-----TIKINSTONNIIVDPDLATA 591
      | | | | | | | | | | | | | | | | | | | | | | | |
470 ---RGYGSATHSQSVTTISWEGVFERG--GEVYSSEKTWISLIRAKVNFMTA 323
      | | | | | | | | | | | | | | | | | | | | | | | |
592 TPLDGVVFASSLINAGTILINCK IGITGANNKTIENIGSKTYLNSGVAVINELVIG 650
      | | | | | | | | | | | | | | | | | | | | | | | |
324 -----KQGVYASQNHAWFQWTF-----GVYQAGSTAMNWTLSLSQV 365
      | | | | | | | | | | | | | | | | | | | | | | | |
651 NDGAVDFAHDYLLIRITNAA---QGRKIIIPVVRNSGITLAAGTNE-----GSAIN 699
      | | | | | | | | | | | | | | | | | | | | | | | |
466 KSNKRGSNRGVADGAMCGAVVITDAVKKILIPG-3SPDYQDSADATIDAHITTL----- 419
      | | | | | | | | | | | | | | | | | | | | | | | |
700 PLATINFGSKGVNVDI-----VLNAGVAVIATNITITIDANGSFEVFNAG 745
      | | | | | | | | | | | | | | | | | | | | | | | |
420 -----GEPTSPNTVFASSGIVFA--KTFHTSVILPDSFTITGGORG--1P 463
      | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 746 GTNIVSGIVSGQGKKNFTVALENCTYKFLGNATFNNGNTTAAINSTIOLGANYADCAV 805
QY 464 FEDSTPFTPEIYVFEQUTFEKQKNNSIVRAVHYSILILPGVPFN--GGGCIQCDCTN 521
      | | | | | | | | | | | | | | | | | | | | | | | |
Db 806 SADPTGIVEFVNTGFTIVTLNKO--AAIVNALKQIIVSGPVVINEIGNAG-----N 856
      | | | | | | | | | | | | | | | | | | | | | | | |
QY 822 H-----PDAQIFPNTL YNSNGN---LATREKIIKISTUSVAVGSRITIS 563
      | | | | | | | | | | | | | | | | | | | | | | | |
Db 857 HRCVAVTTIAEENSLGAVFLPRGIFPDAGNIMPILIKSTVGNKTKAGFDPVSVVVLG 916
QY 564 TDSISKASLIINVTGATHTVNTGPRIPITLIMQDNEYSFQVSDSGV--ALPGYMWL 620
      | | | | | | | | | | | | | | | | | | | | | | | |
Db 917 VDSVYADGQVI-----GQONNIVGLGLGSDNGLIYVATTLIYAGI 955
      | | | | | | | | | | | | | | | | | | | | | | | |
QY 621 FVNN-----SAGVPSVASTI 635
      | | | | | | | | | | | | | | | | | | | | | | | |
Db 956 STINNNOGTVTLGGVDPNPQTV 978

```

Search completed: January 3, 2002, 09:19:39
 Job time: 270 sec

Database version 4.1
Copyright (c) 1991-2000 Computer Ltd.
M Protein - protein search using sequence

Run on: January 6 2002 09:50:08 - 09:50:20 (00:00:12)
(without alignment)
100.0024 Million cell updates/sec

US-09-722-602-21
1 ASAP1991-1999MAY1999
1999-02-19 09:50:08

Scoring Table:
Gapop 1000 - Gapop 1000

Scoring: 47505 seqs, 14627249 residues

Total number of hits satisfying chosen parameters: 47505

Minimum DB seq length: 6

Maximum DB seq length: 20000000

Post-processing: 10000000

Maximum Match: 1000
Listing first 45 similarities

Database: 1: SPATHEA*
2: SPATHEA*
3: SPATHEA*
4: SPATHEA*
5: SPATHEA*
6: SPATHEA*
7: SPATHEA*
8: SPATHEA*
9: SPATHEA*
10: SPATHEA*
11: SPATHEA*
12: SPATHEA*
13: SPATHEA*
14: SPATHEA*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Prod. No.	Score	Query	Match	Length	ID	ID	Description
1	418.5	12.2	526	2	05679	05679	Stigmatella
2	337.5	9.3	446	2	05679	05679	Stigmatella
3	296.5	8.7	515	10	05679	05679	Stigmatella
4	275.5	8.0	547	10	05679	05679	Stigmatella
5	272	7.9	504	10	05679	05679	Stigmatella
6	266	7.8	515	10	05679	05679	Stigmatella
7	243	7.1	45	4	04098	04098	Stigmatella
8	241.5	7.1	541	10	05679	05679	Stigmatella
9	206.5	6.0	1196	2	05679	05679	Stigmatella
10	200	5.8	1446	2	05679	05679	Stigmatella
11	197	5.9	1441	2	05679	05679	Stigmatella
12	184.5	5.4	559	4	05173	05173	Stigmatella
13	184.5	5.4	559	4	05173	05173	Stigmatella
14	178	5.2	1441	2	05679	05679	Stigmatella
15	163.5	4.8	1441	2	05679	05679	Stigmatella
16	161.5	4.7	1441	2	05679	05679	Stigmatella
17	161.5	4.7	1441	2	05679	05679	Stigmatella
18	161	4.7	1441	2	05679	05679	Stigmatella
19	160	4.7	1441	2	05679	05679	Stigmatella

20	159	4.6	634	4	05679	05679	Stigmatella
21	158	4.6	1323	6	05679	05679	Stigmatella
22	158	4.6	1800	2	05679	05679	Stigmatella
23	155.5	4.5	856	3	05679	05679	Stigmatella
24	152	4.4	1219	2	05679	05679	Stigmatella
25	150.5	4.4	670	2	05679	05679	Stigmatella
26	150	4.4	543	2	05679	05679	Stigmatella
27	149	4.4	2271	2	05679	05679	Stigmatella
28	148.5	4.3	1875	5	05679	05679	Stigmatella
29	148.5	4.3	1875	5	05679	05679	Stigmatella
30	147.5	4.3	667	3	05679	05679	Stigmatella
31	147	4.3	481	6	05679	05679	Stigmatella
32	146	4.3	1617	2	05679	05679	Stigmatella
33	145	4.3	1617	2	05679	05679	Stigmatella
34	145.5	4.2	713	2	05679	05679	Stigmatella
35	145.5	4.2	1040	2	05679	05679	Stigmatella
36	145	4.2	1001	3	05679	05679	Stigmatella
37	144.5	4.2	1275	5	05679	05679	Stigmatella
38	144	4.2	1180	2	05679	05679	Stigmatella
39	142.5	4.2	848	2	05679	05679	Stigmatella
40	141	4.1	13055	5	05679	05679	Stigmatella
41	140.5	4.1	2349	2	05679	05679	Stigmatella
42	140.5	4.1	2383	2	05679	05679	Stigmatella
43	140	4.1	783	1	05679	05679	Stigmatella
44	139.5	4.1	750	2	05679	05679	Stigmatella
45	139.5	4.1	750	2	05679	05679	Stigmatella

ALIGNMENTS

RESULT	ID	PHYLIMINARY	PRT	526 AA
05679	05679	PHYLIMINARY	PRT	526 AA
AC	05679	PHYLIMINARY	PRT	526 AA
DT	01-NOV-1996 (TREMEL, 01, Created)			
DT	01-NOV-1996 (TREMEL, 01, Last sequence update)			
DT	01-JUN-2001 (TREMEL, 17, Last annotation update)			
DE	SIMILARITY TO GALACTOSE OXIDASE FROM TACILUM DENROIDES.			
GN	FBP GENE.			
OS	Stigmatella aurantiaca.			
OC	Bacteria; Proteobacteria; delta subdivision; Myxobacteria;			
OX	Myxococcales; Cytochromes; Cytochromes; Stigmatella.			
RN	NCBI_TaxID=41;			
RP	SEQUENCE FROM N.A.			
KC	STRAIN=DM4/3-1;			
RX	MEDLINE=97113426, PubMed=8952286,			
FA	Stigmatella aurantiaca fruiting body formation is dependent on the			
RT	trbA gene encoding a polypeptide homologous to chitin synthases.			
RI	J. Bacteriol 178:6706-6711(1996)			
DR	EMBL: Z11601; CAA77680.1;			
DR	InterPro: IPR001798; Ketch.			
DR	Plant: P01344; Ketch.			
SO	SEQUENCE 526 AA: 57820 MW: 837146012ZACBD CRC64:			
07	148 ASYTAPO--GICWGPTEPPIVPAANAIEPTSGVIMMSYFNAAGSGSGITITS 205			
08	57 SSQAVAGPISVGFWSPI MSWP-SAIHPIHUSKVM-----PAPPGVGSIP 109			
09	206 SMPPTVSEFVETVTRKMPPTKMDQIVTQCN----PAPVSTVSPSPW 260			
10	110 EMPPLANTL--PIHAPPNICAGHSPLPGRITITGSHVSHVVPVPIATNPSCAM 167			
11	261 DPGMNVAPVQSSVATMSPGVPVITGSSNG-CVFFKMPVYSSPSKRTISLPANRVP 319			
12	168 DVFENRERKWFNTITANGDVLALSGEDGDLNLEPFRVAVATNSMGN----- 220			

[illegible][illegible]

1 SILAVVSLAAATLAAFAASDAAG-WREDEKINLSGVALEAIVNSSLVLEF---DRAT 57
 165 GASP---GGITITSSWDSTGIVSRIVVTKHMFPGISLSDGNGQIVTGGNDA--- 247
 58 GQPLKINSGTWGALMDITSVRLSVLT---LSPCAGALLSNITVMSMGCTPGG 114
 248 -----KRTSLVSSSSSWIPGHP-----MVAAGVSSAIMEGCVETI 286
 115 GVAAPPCGQAIIRIFPCAS---PSIDGCTLFEDPATVHLLEKRWYSSVRIEDGSIIM 171
 247 GGSWSGVFPKNGEYSP-----SSKWTSTINAKVNPMITADKQGLYRSDN 333
 172 GGS---HVLTPYVNDPANSFEPTPSKQIIPDSAFLESLPAN---LPR-- 216
 444 HMTLPGWKEGCVLAHSLAMKVTYSSTVVSATKEGSPGVAPT-----AMCG 384
 217 ---ATAIPGQIVIVANNOSIIV-----DIE-----KNETILPDIPNGVRYTNPIDG 261
 385 NAWV-----DAVKKILTFGSPDYODSDATDAHITLIGEPGTS-----PNTVFA 431
 262 SALLPLSPDPTF-EVLVCGGS---TADTSLPSTLSQHPATISQSRITTIPEGTKA 316
 442 S---NGLYFAPIHTSVLPDSTFTLGGQPRG-----PEPDSI---PVTPELYV 477
 417 GWOVEHMLEAKMPELVHYNGOILLNGAGTFAALSAVADPVGNNAHPVLIPSLYI 376
 478 PRDITFYKQNPN-----SIVKAYHSISILLPGRVNGG-----GLGQDCTNH 522
 477 PCK-PGKRISNAGMPTTIPKMHSTVTLTGQNPFTGQNPMPNPTPGIPEPSE 435
 523 PAAITITNTLNSNGLAIPKLTETSTGSKVQGR-----ITISTSSISK-ASTIRY 576
 446 LREILDPPEMFESKALLMP-----EKLKQGVIVITLPSLKA-SKVALMEL 488
 577 GIVATVNTDQRPILPTLNNG-NSYSPVPSDGVAPGVMFLF-----VM 623
 489 GFSHAFHSARLVFMSSISADKKSILFTAPNGKRVTPGPVAVFILLIVTSGPERVM 548
 624 NSACVP 629
 447 MSINP 554
 14
 501772 PRELIMINARY; PRI: 559 AA
 01-N-V-1996 (TREMBLrel. 01, Created)
 01-N-V-1996 (TREMBLrel. 01, Last sequence update)
 01-N-V-1996 (TREMBLrel. 08, Last annotation update)
 GLYOXAL OXIDASE (GLX1) PRECURSOR (GLX1).
 Phloroglucinol chrysoportin.
 Eukaryotic, Fungi, Basidiomycota, Hymencomycetidae;
 Agaricomycetes; Corticiaceae; Phanerochaete.
 NBL_TaxID: 5309;
 SEQUENCE FROM N.A.
 Kristen P., Witek C., Vanden Wymelenberg A., Cullen D.;
 J. Bacteriol. 0:0-0(0).
 [2]
 SEQUENCE OF 1-529 FROM N.A.
 MEDLINE: 9348282; PubMed: 8346264;
 Kristen P., Cullen D.;
 "Cloning and characterization of cDNA encoding glyoxal oxidase, a
 H₂O₂-producing enzyme from the lignin-degrading basidiomycete
 Phanerochaete chrysosporium."
 Proc. Natl. Acad. Sci. U.S.A. 90:7411-7412(1993).
 EMBL: 147286; AAA87594.1;
 EMBL: 147991; AAA3747.1;
 Signal: Peroxidase.
 RT
 SIGNAL 22 POTENTIAL.
 RT
 CHAIN 28 559 POTENTIAL.

SQ SEQUENCE 559 AA; 59167 MW; 9245FA19E198B65 GR64;
 Query Match 5.4%; Score 183.5; DB 4; Length 559;
 Best local similarity 21.8%; Pred. No. 90-05;
 Matches 132; Conservative 84; Mismatches 221; Indels 169; Gaps 30;
 139 SLAEINVOASSTYAP---QVGLCHWQPIIDLPYVAAAALPISGRVIMSSYNDAF 194
 3 SILAVVSLAAATLAAFAASDAAG-WREDEKINLSGVALEAIVNSSLVLEF---DRAT 57
 195 GASP---GGITITSSWDSTGIVSRIVVTKHMFPGISLSDGNGQIVTGGNDA--- 247
 58 GQPLKINSGTWGALMDITSVRLSVLT---LSPCAGALLSNITVMSMGCTPGG 114
 248 -----KRTSLVSSSSSWIPGHP-----MVAAGVSSAIMEGCVETI 286
 115 GVAAPPCGQAIIRIFPCAS---PSIDGCTLFEDPATVHLLEKRWYSSVRIEDGSIIM 171
 247 GGSWSGVFPKNGEYSP-----SSKWTSTINAKVNPMITADKQGLYRSDN 333
 172 GGS---HVLTPYVNDPANSFEPTPSKQIIPDSAFLESLPAN---LPR-- 216
 444 HMTLPGWKEGCVLAHSLAMKVTYSSTVVSATKEGSPGVAPT-----AMCG 384
 217 ---ATAIPGQIVIVANNOSIIV-----DIE-----KNETILPDIPNGVRYTNPIDG 261
 385 NAWV-----DAVKKILTFGSPDYODSDATDAHITLIGEPGTS-----PNTVFA 431
 262 SALLPLSPDPTF-EVLVCGGS---TADTSLPSTLSQHPATISQSRITTIPEGTKA 316
 442 S---NGLYFAPIHTSVLPDSTFTLGGQPRG-----PEPDSI---PVTPELYV 477
 417 GWOVEHMLEAKMPELVHYNGOILLNGAGTFAALSAVADPVGNNAHPVLIPSLYI 376
 478 PRDITFYKQNPN-----SIVKAYHSISILLPGRVNGG-----GLGQDCTNH 522
 477 PCK-PGKRISNAGMPTTIPKMHSTVTLTGQNPFTGQNPMPNPTPGIPEPSE 435
 523 PAAITITNTLNSNGLAIPKLTETSTGSKVQGR-----ITISTSSISK-ASTIRY 576
 446 LREILDPPEMFESKALLMP-----EKLKQGVIVITLPSLKA-SKVALMEL 488
 577 GIVATVNTDQRPILPTLNNG-NSYSPVPSDGVAPGVMFLF-----VM 623
 489 GFSHAFHSARLVFMSSISADKKSILFTAPNGKRVTPGPVAVFILLIVTSGPERVM 548
 624 NSACVP 629
 549 MGSNP 554
 RESULT 15
 09CFIL1 PRELIMINARY; PRI: 1441 AA.
 AC 09CFIL1;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE UNKNOWN PROTEIN.
 GN VOIC.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Lactococcus
 OX NBL_TaxID-1360;
 RN SEQUENCE FROM N.A.
 RC STRAIN-111403;
 RA Holton A., Winkler P., Mauger S., Jallion O., Malame K.,
 Weissbach J., Ehrlich S.D., Sorokin A.;
 "The complete genome sequence of the lactic acid bacterium Lactococcus
 lactis."
 Genome Res. 0:0-0(2001).
 RT

EMBL: A006376; AK0556.1; -
 InterPro: IPRO01899; Gran_pos_anchor;
 Prosite: PS00443; GRAM_pos_anchor; UNKNOWN_1;
 Complete proteome;
 SOURCE 1441 AA 14246 MW: 401446501815EA CRC64:

Very Match 5.28; Score 178; LR 2; Length 1441;
 best local similarity 22.0%; Prod. No. 9.00077;
 Matches 158; Conservative 22; Mismatches 212; Indels 266; Gaps 40;

```

87 VYSSDGIWMS--PVASISWEADSTI---KYS---NEELPARYV-----RLVAT 129
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
818 VYQYVTRITRTVTVETSTALANAGALZATETIELTVAITAFASIVLNDVNNKAVS 877
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
130 IFANG-----QFMSIAETNVQASSYTAPO---PGLQWGPIDLPVPAALAE 177
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
878 ALANVIVNEVVAIPIITIKDSIIIDEGSTWIAADNFTGISSTIGVLTINDVTVGTVD 537
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
124 PTSGEVLMSSEYENLAFSGPQITTSK--DPSGIVSDRIVATKDMFGQISMDG- 235
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
938 IS-----PQSSSVIYIIDPTIGASISSVANITVNDSSNTINPTDGT 980
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
246 -----NGLVYTONAKKISLYDSSSDSWIPQPMQVAKGYQSATM--SD-GRVF 284
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
981 GGGNTINPTDENTIVLPGAGGR---LQIADQVYVIG DQFLSPGDSGVLTLESDGKVD 1035
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
285 TTGGSMS--GVFFANGGVY-----SPSKTWTG-----LQNAKVNM 320
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1036 RPDISYVNVPGIVVPEGTIHLFGIVINPGSGSVIVPGPQKIGIDDTLINPNSPVVP- 1094
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
321 LTADKQGYRQGNHAMEFGWKKOSVFCATSTAM--NWYYT----- 359
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1095 -----GDN-----GSVTLPGQGIASIPNGNITLFGIVVDPDGIHLPGSD 1135
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
360 -----SSGVYSAGKESNECAVEDACGNAVATDAVKKILLFGGSP 403
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1136 IVNEGSTITLPGQKRTSTDE--DCKIKPRPTLPG -LNGSVTLPGQGVITPQGIT 1189
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
404 IYO-----DSATIDAH-----IT-----GPDTSPTNVEASNG----- 434
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1190 NVPGDSVVPPEGT--VHLQELVNEFMIIITLPGQDKGTGTNDGKVRKNGPSISNDGS 1247
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
435 --LVEARFHT-----SVLPGDSIFITGQRRGIPREDSTVFTPEIYPEQ-- 480
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1248 LILPGGGIVITPGCTINVPDGSVVDPDQIVHLPG--GVNPDGI-----ITLPGQDG 1298
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 -----DPEYKQNPSTVEAVHSISL-----LLEPRVFNQSGSGDGTINHPDQIFT 529
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1299 KTIGIEVNIKFNPSVSNIHPSIILPGGJIVTNGGTYNVPGIVDP-----DGTIHL 1353
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
530 PN-VLYNSNMLAKRTFLPSTCEVKKCKK- -IISGSSISKASITRYGTAIHIVNID 586
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1354 PNDGVINPDIVSGKSNSTIFRNGGTSSTANTLSSNGPS-----TLNSN 1401
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
587 QRRIPIILNNGSNYSFQVDSGVALPGYMLLVNNSA-----GVPSVASTIVTQ 639
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1402 NTLIPKI---GDSS-----MNSAEALAGISALIGVMTLLE 1434
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Job completed: January 3, 2002, 09:19:01
 Job time: 252 sec



Feature: linear
Molecule type: protein
OS: us-09-722-602-21

Query Match: 100% Score 141 2525 Length 157A
Best Local Similarity: 22.4% Pos: 100%
Matches: 132 Conservative: 64 Mismatches: 189 Indels: 186 Gaps: 44

14 WAVI HSA:NMTH:KALISRETTWITLNTT...
153 WAGNIAHAIN:NMTH:TAHTREAHTVQSSVNGTANAKTITTTTAASTAN 245
65 VNGJSMH:NMTH:NMTHVYSQATNMISVASTMA...
250 GAST...
100 PARYLVALLH:NMTH:SLAFVQASST...
209...
171 PAAA...
16 FEALNEN...
200 GILLIS...
193 POSV...
243 GONAKK...
456 G...
292 JAVERN...
602 NATHIE...
448 ANPSTAM...
920...
395...
607 HRRH...
439 RFE...
608...
609...
610...
611...
612...
613...
614...
615...
616...
617...
618...
619...
620...
621...
622...
623...
624...
625...
626...
627...
628...
629...
630...
631...
632...
633...
634...
635...
636...
637...
638...
639...
640...
641...
642...
643...
644...
645...
646...
647...
648...
649...
650...
651...
652...
653...
654...
655...
656...
657...
658...
659...
660...
661...
662...
663...
664...
665...
666...
667...
668...
669...
670...
671...
672...
673...
674...
675...
676...
677...
678...
679...
680...
681...
682...
683...
684...
685...
686...
687...
688...
689...
690...
691...
692...
693...
694...
695...
696...
697...
698...
699...
700...
701...
702...
703...
704...
705...
706...
707...
708...
709...
710...
711...
712...
713...
714...
715...
716...
717...
718...
719...
720...
721...
722...
723...
724...
725...
726...
727...
728...
729...
730...
731...
732...
733...
734...
735...
736...
737...
738...
739...
740...
741...
742...
743...
744...
745...
746...
747...
748...
749...
750...
751...
752...
753...
754...
755...
756...
757...
758...
759...
760...
761...
762...
763...
764...
765...
766...
767...
768...
769...
770...
771...
772...
773...
774...
775...
776...
777...
778...
779...
780...
781...
782...
783...
784...
785...
786...
787...
788...
789...
790...
791...
792...
793...
794...
795...
796...
797...
798...
799...
800...
801...
802...
803...
804...
805...
806...
807...
808...
809...
810...
811...
812...
813...
814...
815...
816...
817...
818...
819...
820...
821...
822...
823...
824...
825...
826...
827...
828...
829...
830...
831...
832...
833...
834...
835...
836...
837...
838...
839...
840...
841...
842...
843...
844...
845...
846...
847...
848...
849...
850...
851...
852...
853...
854...
855...
856...
857...
858...
859...
860...
861...
862...
863...
864...
865...
866...
867...
868...
869...
870...
871...
872...
873...
874...
875...
876...
877...
878...
879...
880...
881...
882...
883...
884...
885...
886...
887...
888...
889...
890...
891...
892...
893...
894...
895...
896...
897...
898...
899...
900...
901...
902...
903...
904...
905...
906...
907...
908...
909...
910...
911...
912...
913...
914...
915...
916...
917...
918...
919...
920...
921...
922...
923...
924...
925...
926...
927...
928...
929...
930...
931...
932...
933...
934...
935...
936...
937...
938...
939...
940...
941...
942...
943...
944...
945...
946...
947...
948...
949...
950...
951...
952...
953...
954...
955...
956...
957...
958...
959...
960...
961...
962...
963...
964...
965...
966...
967...
968...
969...
970...
971...
972...
973...
974...
975...
976...
977...
978...
979...
980...
981...
982...
983...
984...
985...
986...
987...
988...
989...
990...
991...
992...
993...
994...
995...
996...
997...
998...
999...
1000...

Query Match: 3.8% Score 131 DB 1: Length 846;
Best Local Similarity: 19.0% Pos: 0.016;
Matches: 132 Conservative: 64 Mismatches: 210 Indels: 258 Gaps: 30;

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/731,157A
FILING DATE: 19910509
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 90209962
FILING DATE: 18-APR-1990
APPLICANT/AGENT INFORMATION:
NAME: RAE-VENTER PH.D., BARBARA
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: GPRO-027/0005
TELEPHONE: 415-494-7622
TELEFAX: 415-457-0663
TELEX: 380816 COOLEY PA
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 846 amino acids
TYPE: AMINO ACID
FEATURES:
1 PEPTIDE: linear
MOLECULE TYPE: protein
ORGANISM: Escherichia coli
PUBLICATION INFORMATION:
AUTHORS: Schumacher, G
AUTHORS: Schumacher, D
AUTHORS: Haug, H
AUTHORS: Hockel, P
AUTHORS: Hockel, A
TITLE: Fucosylid acylase from E. coli: unique
JOURNAL: Nucleic Acids Res.
VOLUME: 14
PAGES: 5713-5727
DATE: 1986
US-07-731-157A-5

107...
132...
144...
447...
507...
219...
553...
278...
601...

EARLIER FILING DATE: 1998-07-02
 EARLIER APPLICATION NUMBER: 60/094,354
 EARLIER FILING DATE: 1998-07-28
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: FASTSO for windows version 3.0
 SEQ ID NO: 2
 LENGTH: 915
 TYPE: PRT
 ORGANISM: Bacillus thuringiensis
 FEATURE:
 NAME/KEY:
 DEFINITION: (1)-(915)
 OTHER INFORMATION:
 US-09-722-602-21

Query Match 3.8% Score 130 DB 4 Length 915
 Post Local Similarity 19.1% Pred. No. 0.018
 Matches 120 Conservative 62 Mismatches 169 Indels 254 Gaps 33

192 DAFG-----GSP--GRTTISGSPGSDRYVIR----- 223
 42 DSGAASAPLIPDSJLVIVSGW TRVSADYILDSGNEWLYPONSUYPNE 99
 224 -- HMDGRTIS--MGNDIIV-- 256
 100 KRAETAAKAVSNAYLAAAGYVALSGLLGLDASQIVHDIANKULVTSKASL 159
 257 -- -- -- -- -- 279
 156 GQNVAVIAGIPHTPGSPGSDRYVIR----- 279
 280 -- -- -- -- -- 279
 220 PAVSWAVMPTVGLTAMTGLV RGLALDSVATILVILSHDQVSLISLQ 276
 325 KQI-----LYSEN----- 345
 277 -- -- -- -- -- 336
 446 PAAQSIAM-----NMY--ISGRIVSAVSGSHGVATACCAVAM 388
 437 TRIVPILASJHWEAVRRLNMTYRLEVTAGSLKIA-----VDPA--TALA 385
 389 YAAVAKKILLEGSDYQSDA--ETDAI-----ITLGPSTSPNVASN 433
 386 PHTGKMTV-----DIAETDAGWNSDKHITPKNIEDVYEDVDYDFSIDPNSGKMK 439
 434 GYFAPPIHISVLEEDSTITGSGRTTISGSPGSDRYVIR-----PEI 475
 440 GYLAIT-FRTKGPNNVIGISLKGKLTHTYQLKPYASNSVDLDTQLQWNGYLPKN 498
 476 YVPPPTPEYENRISVAAHNSILSPGAVFNSGGLGGLINH+DAQI-----F 528
 466 YVPPPTPEYENRISVAAHNSILSPGAVFNSGGLGGLINH+DAQI-----F 528
 529 PPNLYNSN-----SN-TAIPKRTTISGSPGSDRYVIR-----GRTTISGSDS 567
 554 YVPPPTPEYENRISVAAHNSILSPGAVFNSGGLGGLINH+DAQI-----F 528
 554 YVPPPTPEYENRISVAAHNSILSPGAVFNSGGLGGLINH+DAQI-----F 528
 554 YVPPPTPEYENRISVAAHNSILSPGAVFNSGGLGGLINH+DAQI-----F 528
 612 LCKRCHKAASGLAHN 628

APPLICANT: MITTA, Masanori
 APPLICANT: ASADA, Kiyozo
 APPLICANT: TSUNASAWA, Susumu
 APPLICANT: KATO, Ikunoshin
 TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Browdy and Neimark
 STREET: 419 Seventh Street N.W., Ste. 300
 CITY: Washington
 STATE: D.C.
 COUNTRY: United States of America
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patcom Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/722-602-21
 FILING DATE: 20-MAY-1998
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US/95/03253
 FILING DATE: 07-NOV-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 323285/1995
 FILING DATE: 12-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Browdy, Roger L.
 REGISTRATION NUMBER: 25,618
 PREFERENCE/DEPT NUMBER: TAKAYURA=1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-5197
 TELEFAX: (202) 737-3528
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 522 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FEATURE:
 OTHER INFORMATION: /note= Xaa at position 428 is Gly or Val.
 US-09-722-602-21-3

Query Match 3.8% Score 130 DB 4 Length 522
 Post Local Similarity 19.1% Pred. No. 0.0091
 Matches 90 Conservative 57 Mismatches 151 Indels 172 Gaps 20

90 SSGTNNMGSPVASCWFDSTRTKYSNFTFRPARYVLAIVTAAGUPWTSIAEINFGAS 149
 149 SSGTNNMGSPVASCWFDSTRTKYSNFTFRPARYVLAIVTAAGUPWTSIAEINFGAS 149
 150 SYTAPQPLGKMGPTIDPIV-----AAAIEPTSGVLMSSVRNDAPGSPG 199
 192 KYVYTSRSGF GTATPRKPFVATKMIIAAVASGISMGGT-----NAYTAAQ 244
 200 GTIT-----SSND--STGIVSDITVIVLKHDMFC-----PG 240
 245 TSMATIHVATIAILLVHMTFTIRKVTAL-ETADIVKFEELADIAYGAGRVANAYA 302
 231 ISMDNGQIVVTG--GNDARKTSIYDSSDSWIPGPDQVAVGYSSATMSGRVTTIG 288
 303 INYDNVAKLVFTGVANKSGTHQFVIGASFY-----TATL----- 339
 289 SMSGVFEKNEPVSPSKTWTSLPNAKVNPLTA-----DKGLYRSNHWALFGWKKS 344
 340 YWDAANSUDDLIYDFNGN-----GVDSYTAAYGFEKVGYYNPDTGWTI----- 385
 345 VFAAGPSTAMMYTSSGSDVKSAGKRSNKGVAAPFANGAANVDAVAKSLITGGSD 404


```

440 -----LYNG----- AIYVIRIVIMANGVMAIVNGLINWYRYV1000S 384
441
442
443
444
445 VESAKPUSNIGVARIAM INAVYVAVKILITFUSIYJUSIAF---LEAH----- 416
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

Status: Completed: January 6, 2002: 15:17:05
 Run Time: 187 sec

•
•
•